A Perturbational Approach for Approximating Heterogeneous Agent Models*

Anmol Bhandari
University of Minnesota

Thomas Bourany
University of Chicago

David Evans
University of Oregon

Mikhail Golosov
University of Chicago

September 2023

Abstract

We develop a perturbational technique to approximate equilibria of a wide class of discrete-time dynamic stochastic general equilibrium heterogeneous-agent models with complex state spaces, including multi-dimensional distributions of endogenous variables. We show that approximating policy functions and stochastic process that governs the distributional state to any order is equivalent to solving small systems of linear equations that characterize values of certain directional derivatives. We analytically derive the coefficients of these linear systems and show that they satisfy simple recursive relations, making their numerical implementation quick and efficient. Compared to existing state-of-the-art techniques, our method is faster in constructing first-order approximations and extends to higher orders, capturing the effects of risk that are ignored by many current methods. We illustrate how to apply our method to a broad set of questions such as impacts of first- and second-moment shocks, welfare effect of macroeconomic risk and stabilization policies, endogenous household portfolio formation, and transition dynamics in heterogeneous agent general equilibrium settings.

*We thank Adrien Auclert, Ben Moll, participants of SITE conference, and seminar participants at Bank of Canada, Bank of Japan, Bocconi, EIEF, EUI, Minneapolis Fed, University of Chicago, University of Oregon, NBER Summer Institute, and the NY Fed for helpful comments. Bhandari, Evans, and Golosov thank the NSF for support (grant #36354.00.00.00).

1 Introduction

We develop a numerical method to approximate equilibrium dynamics of a large class of discrete-time heterogeneous agent (HA) models that feature aggregate and idiosyncratic shocks, and occasionally binding borrowing constraints. Our method is based on approximation techniques that scale aggregate shocks and consider Taylor expansions of equilibrium conditions with respect to that scaling parameter. Our method can be used to quickly compute equilibrium effects of higher-moment shocks, welfare effects of risk and macroeconomic stabilization policies, general equilibrium portfolio problems, and transition dynamics to a new steady state in environments with rich heterogeneity.

Our approach intentionally focuses on maintaining the computational speed, flexibility, and ease-of-use of traditional perturbational techniques commonly employed to solve and estimate representative agent dynamic stochastic general equilibrium models, such as those implemented with the DYNARE software package.[1] We use recursive representations of equilibria and approximate equilibrium dynamics around the steady state of the model in which aggregate shocks are switched off. We show that it is possible to analytically derive expressions for all objects that are used to construct equilibrium approximations. These expressions are represented by a small-dimensional linear system of equations regardless of the dimensionality of the state in the recursive equilibrium representation (which is usually infinite-dimensional in HA economies). In our implementation, the construction of these linear systems requires just two inputs: (i) the standard outputs from commonly used routines for calculating the steady state (i.e., the invariant distribution) of HA economies without aggregate shocks, and (ii) the functional forms defining equilibrium conditions. In addition to preserving the spirit and convenience of DYNARE-like algorithms, our method is faster than existing state-of-the-art techniques to obtain the first-order approximations of HA economies, and extends easily to higher orders.

Our approximation approach is based on two insights: (i) that traditional perturbational techniques can be reformulated as solving a small-dimensional linear system that characterize values of certain directional derivatives; and (ii) that this reformulation is particularly useful for HA economies, which have a high-dimensional state and feature kinks in policy functions, for example, due to the presence of occasionally binding borrowing constraints. We describe each of these insights in turn.

Policy functions in the recursive representation depend on exogenous (idiosyncratic and aggregate) shocks and endogenous (typically, multi-dimensional) state variables. Traditional application of perturbational techniques (e.g., Schmitt-Grohé and Uribe (2004)) requires finding derivatives of policy functions with respect to all its arguments, evaluated at the steady state. These derivatives solve non-linear matrix equations whose size and complexity scales with the dimension of the state. Using these derivatives, one can construct equilibrium responses of desired order of approximation to aggregate

[1]See Judd (1998) and Schmitt-Grohé and Uribe (2004) for an introduction to such methods.

shocks.

Our first contribution is to demonstrate a reformulation of the perturbational approach. Specifically, we illustrate that aggregate equilibrium responses can be characterized using *directional derivatives*—values of policy function derivatives evaluated in suitably chosen directions. This reformulation offers several advantages. Firstly, the directional derivative of any policy function remains scalar, ensuring its dimension is one regardless of the state variable's dimensionality. This is particularly salient in HA contexts, where the endogenous state's dimensionality is vast, yet key equilibrium objects are summarized by a limited set of variables. Secondly, the directional derivatives we employ have intuitive economic interpretations. For instance, the first-order approximation is derived from a sequence of directions that align with the state's equilibrium path following a one-time, unanticipated aggregate disruption (commonly known as an "MIT shock"). Second-order approximations use directional derivatives reflecting the state's trajectory due to compounded MIT shocks over different periods and a direction capturing precautionary motives due to the presence of risk. Lastly, all directional derivatives solve linear equations with coefficients that can be deduced analytically. Even in representative agent (RA) scenarios, our method presents benefits. For example, it circumvents the need to tackle nonlinear equations or incorporate further refinements – such as selecting the stable root of matrix quadratic equations or resorting to pruning – which traditional perturbational techniques require to rule out approximations with explosive paths.

Applying perturbational techniques to HA settings introduces distinct challenges. Firstly, the presence of idiosyncratic risk and incomplete markets means that equilibrium dynamics typically hinge on an infinite-dimensional distributional state. Our formulation readily adjusts to this and we show that both individual and aggregate responses can be expressed using only a small set of directional derivatives that are constructed using convenient linear operators. Secondly, traditional perturbational methods rely on the assumption that policy functions are sufficiently differentiable at the steady state. This assumption is violated in many staple HA models due to the presence of occasionally binding borrowing constraints, which induce kinks in policy functions. These kinks are endogenous in the sense that they themselves depend on the state of the economy. We show that such kinks can be explicitly incorporated into analysis using *generalized functions*.[2] Leveraging these generalized functions allows us to characterize the effects of kinks in policy functions and mass points in the invariant distribution on equilibrium responses to aggregate shocks.

To apply our approach in HA economies, we first derive the exact expressions for the linear opera-

[2] The generalized functions, (e.g, functions that include delta-function components) are also commonly known as distributions. Mathematically, these functions are not related to distributions in the sense of probability theory. Thus, we reserve the word "distribution" for the latter (e.g., the invariant distribution of asset holdings in HA economies) and refer to the former as "generalized functions". The reader can consult Kanwal (1998) for a good introduction to generalized functions.

tors that characterize the theoretical values of the required directional derivatives and then propose a discretization scheme to implement them numerically. This sequence ensures that as the approximation grid shrinks in size, our numerical solutions consistently converge to the true solution. Notably, such convergence would not be guaranteed if we followed the traditional technique of first discretizing the transition probability kernel that characterizes the steady state invariant distribution and then applying perturbational techniques. In particular, we show that the most popular method to discretize the transitional probability kernel – the so-called histogram method – cannot be used to study equilibrium approximations beyond the first order as it fails to correctly capture non-linearities in the laws of motion in the aggregate distribution of HA economies.

In Sections 2–4, we focus on the most canonical formulation of the approximation problem – the economy starts near its steady state, aggregate shocks are homoskedastic, and all equilibrium variables are uniquely determined in the deterministic economy. We first develop our analytical techniques in RA settings (Section 2), and then extend them to the HA settings (Section 3). In Section 4, we discuss how outputs from off-the-shelf algorithms used to compute the steady state of the deterministic model can be used to construct objects needed for first- and second-order approximations to aggregate shocks. In Section 5, we consider three extensions that do not fit into our baseline framework: models with transition dynamics to a steady state, models with stochastic volatility, and portfolio problems.

Models with transition dynamics emerge when initial conditions are different from their long-run values. Such problems arise naturally when one studies the effects of permanent shocks or policy changes. We show that the same linear operators that we derived to approximate around the non-stochastic steady state also characterize the equilibrium path from a given initial condition to the long-run ergodic distribution.

Models featuring stochastic volatility are prevalent in many settings that explore the implications of changes in uncertainty and risk premia. The conventional approach to incorporating time-varying risk can be cumbersome even in simple settings. The reason is that when employing standard perturbational techniques, both the level and volatility of innovations to aggregate shocks are perturbed, meaning that the effects of volatility shocks only manifest at the third order. We proceed differently: by perturbing only the level of the shock, we can describe the impact of uncertainty shocks using just second-order expansions. Furthermore, our specification of uncertainty shocks combined with the directional derivatives approach retains tractability even in rich HA settings. In fact, we characterize equilibrium responses to uncertainty with the same linear operators we previously constructed for the first and second-order approximations in a homoskedastic economy. This suggests that approximating models with stochastic volatility is as time-efficient as doing so in the homoskedastic settings.

In the last extension of Section 5, we turn attention to portfolio problems. Macroeconomic environments with portfolio problems–models in which agents can invest in assets with different risk characteristics–are commonplace. This encompasses settings where agents can both borrow and lend among themselves and also invest in risky capital. However, these models pose challenges for perturbational methods. In a deterministic economy, all assets are risk-free, making portfolio allocations indeterminate. Moreover, while portfolio choices hinge on the model's second-order properties, like risk premiums or covariances of equilibrium variables, even the model's first-order dynamics are influenced by these choices.

The HA literature predominantly follows two paths. The first essentially sidesteps the household portfolio problem, positing a mutual fund that manages a unified portfolio for all households. The second, meanwhile, disregards the risk attributes of assets and assumes that portfolio allocations are determined entirely by other forces, such as differences in transaction costs of trading various assets.[3] We augment this literature by developing a general approach that incorporates risk considerations. Our method simultaneously solves for the optimal portfolios of all agents, second-order risk premium of these assets, and equilibrium approximation for aggregate responses that depend on those portfolios.

In Section 6, we place our method in the context of the large literature that approximates HA models with aggregate shocks. In Section 7, we use a calibrated economy in the spirit of Krusell and Smith (1998) to compare the precision and speed of our method to existing techniques and then study several applications that demonstrate usefulness of our method. Proofs for all the main statements are in relegated to an online appendix.

2 Perturbational approximations with directional derivatives

In this section, we show that standard perturbational techniques to approximate equilibria in dynamic stochastic macroeconomic models can be reformulated, to any order of approximation, as a problem of finding sequences of directional derivatives that solve linear systems of equations. It will be helpful to present our key ideas first in the simplest, representative agent (RA) settings as most of the insights carry over directly to richer HA models.[4]

Consider the problem of finding solution to a one sector neoclassical growth model, in which agents have preferences $\frac{1}{1-\gamma}\mathbb{E}_0\sum_{t=0}^{\infty}C_t^{1-\gamma}$, technology is Cobb-Douglas $\exp\left(\Theta_t\right)K_t^{\alpha}$, productivity Θ_t follows an AR(1) process

$$\Theta_t = \rho_\Theta\Theta_{t-1} + \mathcal{E}_t, \tag{1}$$

where $\mathcal{E}_t$ is a mean-zero random variable drawn independently across time from a distribution with bounded support, $|\rho_\Theta| < 1$, initial capital stock K_{-1} is given and δ denotes depreciation rate. The

[3] See Gornemann et al. (2021) and several papers that follow them for the mutual fund trick and the large literature on two-asset HANK pioneered by Kaplan et al. (2018) for asset-specific transaction costs.

[4] We thank Ben Moll for the suggestion to use the RA model to explain the key insights of our method.

optimal allocations in this economy can be characterized by three equations as follows

$$C_t^{-\gamma} - \beta \mathbb{E}_t \lambda_{t+1} = 0, \tag{2a}$$

$$\lambda_t - \left(1 + \alpha \exp\left(\Theta_t\right) K_t^{\alpha-1}\right) C_t^{-\gamma} = 0, \tag{2b}$$

$$C_t + K_t - \Theta_t K_{t-1}^{\alpha} - (1-\delta) K_{t-1} = 0. \tag{2c}$$

Let $X_t = [K_t, C_t, \lambda_t]^{\mathrm{T}}$ be the vector of endogenous variables. We are interested in finding approximations to stochastic sequence $\{X_t(\mathcal{E}^t)\}_{t,\mathcal{E}^t}$ that solves equation (2) given stochastic process (1) and the initial condition K_{-1}.

We can express the system of equations (2) more generally. Let $Y_t := [\Theta_t, \mathsf{P}X_{t-1}, X_t, \mathbb{E}_t X_{t+1}]^{\mathrm{T}}$ and write equation (2) as

$$G(Y_t) = 0, \tag{3}$$

where mapping G is explicitly defined by (2) and P is a selection matrix for the pre-determined endogenous variables.[5] Vast majority of DSGE models can be represented in this form, with vector Y_t consisting of exogenous variables, pre-determined endogenous variables $\mathsf{P}X_{t-1}$, current period endogenous variables X_t and their expectations $\mathbb{E}_t X_{t+1}$. Relative to the canonical formulation of such problems given in Schmitt-Grohé and Uribe (2004), we introduced an auxiliary variable λ_t to ensure that expectations are linear in X_{t+1}. This can be done without loss of generality and simplifies analysis.

The perturbational approach approximates the solution to this problem as follows. First, it perturbs the stochastic process (1) by scaling exogenous shocks $\mathcal{E}$ by scalar $\sigma \geq 0$:

$$\Theta_t = \rho_\Theta \Theta_{t-1} + \sigma \mathcal{E}_t. \tag{4}$$

We refer to the economy with $\sigma = 0$ as the *deterministic economy*. Second, the problem is written recursively. In our example, the state variable is $Z = [\Theta, K]$. Using bars to denote policy functions, the recursive formulation of this problem can be written as

$$G\left(\overline{Y}(Z;\sigma)\right) = 0 \text{ for all } Z, \sigma \tag{5}$$

where

$$\overline{Y}(Z;\sigma) = \left[\Theta, K, \overline{X}(Z;\sigma), \mathbb{E}_{\mathcal{E}} \overline{X}\left(\rho_\Theta \Theta + \sigma \mathcal{E}, \overline{K}(Z;\sigma);\sigma\right)\right]^{\mathrm{T}}. \tag{6}$$

[5] Throughout, we assume that G depends only on the subset of Y_t for which this dependence is no-trivial. For example, equation (2) defines a mapping $G : \mathbb{R}^6 \to \mathbb{R}^3$, where G is a function of scalars Θ_t, K_{t-1}, $\mathbb{E}_t \lambda_{t+1}$ and a three dimensional vector X_t. For our exposition, we keep convention that Θ_t and K_t are uni-dimensional variables but extension to multi-dimensional case is straightforward (see the appendix).

Finally, one finds the steady state Z^* of the deterministic economy ($Z^* = [0, K^*]$ in our example) and takes various orders of Taylor expansions of (5) with respect to σ, evaluated at $\sigma = 0$, to approximate the stochastic economy.

To simplify notation, we drop explicit references to σ and Z when $\sigma = 0$ and $Z = Z^*$, that is $\overline{X}(Z)$ and $\overline{X}$ denote $\overline{X}(Z;0)$ and $\overline{X}(Z^*;0)$, respectively. Let $\overline{Z}(Z)$ denote the Law of Motion (LoM) of the state in the deterministic economy, which in our example takes the form $\overline{Z}(Z) = \left[\rho_\Theta \Theta, \overline{K}(Z)\right]$. Finally, we use notation $\overline{X}_\sigma$, $\overline{X}_{\sigma\sigma}$ to denote first two derivatives of $\overline{X}(Z^*, \sigma)$ with respect to σ evaluated at $\sigma = 0$. We refer to these terms as *precautionary motives*.

A standard implementation of perturbational methods (e.g., Schmitt-Grohé and Uribe (2004) or DYNARE) differentiates (5) to find derivatives of policy functions $\overline{X}_Z = \left[\frac{\partial}{\partial \Theta}\overline{X}, \frac{\partial}{\partial K}\overline{X}\right]^{\mathrm{T}}$ and their higher-order generalizations and uses those derivatives to construct equilibrium responses to any sequence of shocks $\mathcal{E}^t$ given initial conditions $K_{-1} = K^*$. Finding $\overline{X}_Z$ requires solving a matrix quadratic equation and picking its stable root.[6] This is easy to do when the dimensionality of Z is small, as it is the case in most RA-DSGE models, but becomes problematic as dimensions of Z grow. The difficulties arise both because it becomes hard to solve such equations and costly to store their solutions.

We present an alternative implementation of the perturbation approach. While being mathematically equivalent to the standard implementation, it does not require solving non-linear equations or storing any matrices that depend on dimensionality of Z. Instead, it collapses the problem of finding the stochastic solution of any order of approximation to solving linear systems of equations where dimensionality does not depend on $\dim Z$. This makes it particularly well-suited for models in which state Z is a large and complicated object, as it is often the case in HA environments.

The pertubational approach implicitly requires that solution is sufficiently stable and policy functions are sufficiently differentiable. We state it here for completeness.[7]

Assumption 1. *Let* $\overline{Z}_t := \underbrace{\overline{Z}(\overline{Z}(....\overline{Z}}_{t\ times}(Z_0)))$.

(a) $\overline{X}(Z;\sigma)$ *is sufficiently differentiable at* $(Z^*, 0)$*;*

(b) $\lim_{t\to\infty} \overline{Z}_t(Z_0) = Z^*$ *for all* Z_0 *in a neighborhood of* Z^**.*

We start by reviewing some basic properties of derivatives. $\overline{X}_Z$ is the Jacobian of $\overline{X}(Z)$ evaluated at $Z = Z^*$, which is a $\dim X \times \dim Z$ matrix that consists of partial derivatives of all policy functions with respect to each variable in Z. For now, let $\overline{X}_Z \cdot \hat{Z}$ be a usual product of matrix $\overline{X}_Z$ and a vector

[6] See Uhlig (2001) and Fernández-Villaverde et al. (2016) handbook chapter to see how to solve the neoclassical growth model using standard pertubational methods.

[7] By "sufficiently differentiable" we mean that policy functions are differentiable at least n times when we consider n^{th} order of approximation. A number of authors (e.g., Blanchard and Kahn (1980) or Jin and Judd (2002)) study sufficient conditions on the primitives of macroeconomic model that ensure that Assumption 1 holds. For our purposes, we simply take this assumption as given.

$\hat{Z}$ of $\dim Z$. Note that we have a relationship

$$\overline{X}_Z \cdot \hat{Z} = \lim_{\alpha \to 0} \frac{1}{\alpha} \left[\overline{X} \left(Z^* + \alpha \hat{Z}; 0 \right) - \overline{X} \left(Z^*; 0 \right) \right], \tag{7}$$

so $\overline{X}_Z \cdot \hat{Z}$ can also be interpreted as a directional derivative, i.e., a measure of how policy functions change if the state Z is perturbed from Z^* in direction $\hat{Z}$. Mathematically, $\overline{X}_Z$ is also the first-order Fréchet derivative of $\overline{X}$ with respect to Z. Fréchet derivatives are helpful because they allow us to generalize (7) to cases when Z is of arbitrary dimensionality. Similarly, $\overline{Z}_Z$ is the Fréchet derivative of $\overline{Z}(Z)$ evaluated at Z^*. In our example it takes the form

$$\overline{Z}_Z = \begin{bmatrix} \rho_\Theta & 0 \\ \frac{\partial}{\partial \Theta} \overline{K} & \frac{\partial}{\partial K} \overline{K} \end{bmatrix}.$$

We are interested in finding approximations to $X_t(\mathcal{E}^t)$ for any $\mathcal{E}^t = (\mathcal{E}_0, ..., \mathcal{E}_t)$. Observe that $X_t(\mathcal{E}^t)$ can be constructed from policy functions as follows

$$X_t(\mathcal{E}^t) = \overline{X} \left(Z_t \left(\mathcal{E}^t; \sigma \right); \sigma \right) \big|_{\sigma=1}, \tag{8}$$

where $Z_t(\mathcal{E}^t, \sigma)$ is defined recursively as $Z_0 = [\sigma \mathcal{E}_0, K^*]$ and

$$Z_t \left(\mathcal{E}^t; \sigma \right) = \left[\rho_\Theta \Theta_{t-1} \left(\mathcal{E}^{t-1}; \sigma \right) + \sigma \mathcal{E}_t, \overline{K} \left(Z_{t-1} \left(\mathcal{E}^{t-1}; \sigma \right); \sigma \right) \right]. \tag{9}$$

Take the first-order Taylor expansion of (8) with respect to σ and evaluate this expression at $\sigma = 0$ to obtain the following characterization of the first-order equilibrium approximation.

Lemma 1. $\overline{X}_\sigma = 0$ *and to the first-order approximation,* X_t *satisfies*

$$X_t \left(\mathcal{E}^t \right) = \overline{X} + \sum_{s=0}^{t} \hat{X}_{t-s} \mathcal{E}_s + O \left(\|\mathcal{E}\|^2 \right),$$

where $\hat{X}_t := \overline{X}_Z \cdot \hat{Z}_t$ *and* $\{\hat{Z}_t\}_t$ *satisfies recursion* $\hat{Z}_0 = [1, 0]^{\mathrm{T}}$ *and* $\hat{Z}_t := \overline{Z}_Z \cdot \hat{Z}_{t-1}$.

Note that σ appears twice in policy functions: it appears as a part of arguments in Z as it scales innovations $\mathcal{E}$, and as a standalone argument. This latter dependence is zero to the first order ($\overline{X}_\sigma = 0$) and so first-order approximations depend only on the first-order responses of policy functions $\overline{X}_Z$ to changes in state $\hat{Z}_t$.[8] Examining the recursion that defines $\{\hat{Z}_t\}_t$ we can see that $\hat{Z}_t$ captures the period-t response of state Z to a one time shock to Θ in period 0. Thus, directional derivatives $\{\hat{X}_t\}_t$ have a natural economic interpretation as the impulse response to an "MIT shock", i.e., to a one-time, unanticipated unit shock to Θ. Lemma 1 is also related to the insight of Boppart et al. (2018) that the first-order approximations can be recovered from impulse responses to MIT shocks.

[8] While the fact that precautionary motives are zero to the first order is well-known (see, e.g., Schmitt-Grohé and Uribe (2004)) it may be helpful to remind the intuition behind this result. The argument σ in $\overline{X}(Z; \sigma)$ captures how the level of risk affects economic decisions. The first-order approximation linearizes economy, and so economic agents behave as if they are risk-neutral and hence risk is irrelevant for them, manifesting in $\overline{X}_\sigma = 0$. This also illustrates that to capture effects of uncertainty on equilibrium variables, one needs to use at least second-order approximations.

To find $\{\hat{X}_t\}_t$ we differentiate (5) with respect to σ and use the observation that in deterministic economy

$$\lim_{\alpha\to 0}\frac{1}{\alpha}\left[\overline{X}\left(\overline{Z}\left(Z^*+\alpha\hat{Z}_t\right)\right)-\overline{X}\left(\overline{Z}\left(Z^*\right)\right)\right]=\overline{X}_Z\cdot\overline{Z}_Z\cdot\hat{Z}_t=\hat{X}_{t+1}.$$

To state our result succinctly, use $K_t = \mathsf{P}X_t$,[9] and let G_Y be the Jacobian of G evaluated at $Y = Y^*$.

Proposition 1. $\left\{\hat{X}_t\right\}_t$ *is the solution to a linear system*

$$\mathsf{G}_Y\hat{Y}_t = 0 \text{ for all } t, \tag{10}$$

where $\hat{Y}_t = \left[\rho_\Theta^t, \mathsf{P}\hat{X}_{t-1}, \hat{X}_t, \hat{X}_{t+1}\right]^{\mathrm{T}}$, $\mathsf{P}\hat{X}_{-1} = 0$, *and* $\lim_{t\to\infty}\hat{X}_t = 0$.

Condition (10) is a linear system of equations that determines $\{\hat{X}_t\}_t$. To construct it numerically, we only need to find matrix G_Y. This is easy to do. The functional form of mapping G is known (see equation (2)) and any automatic differentiation package can obtain the explicit functional form for the gradient of G for arbitrary Y. This equation is then evaluated at $Y = Y^*$ to compute G_Y. Since no numerical differentiation is involved, one obtains the exact value of G_Y using this procedure. The same observation will carry to higher orders, when numerical differentiation becomes less precise and numerically less stable.

Equation (10) is an infinite system of linear equations. Our stability assumption 1b implies a boundary condition $\lim_{t\to\infty}\hat{X}_t = 0$, so the solution can be approximated by truncating this system at some period T, imposing a terminal condition $\hat{X}_t = 0$ for $t > T$ and inverting resulting $(T+1)\times\dim X$ matrix to solve for $\{\hat{X}_t\}_{t=0}^T$.[10] We want to make a couple of observations about this approach. Nowhere in the proofs we relied on the fact that K is uni-dimensional and the same result applies to economies in which endogenous state variable has arbitrary number of dimensions. We also side-stepped the need to solve a matrix quadratic equation or choose its stable root, which would be required under standard implementation of perturbational techniques. These advantages are particularly salient when dimensionality of Z is large, as will get clear in the next section.

Our directional derivative approach extends to arbitrary orders of approximation. We focus on the second order in this paper. Let $\overline{X}_{ZZ}$ be the second-order Fréchet derivative of $\overline{X}$ and $\overline{X}_{ZZ}\cdot(\hat{Z}',\hat{Z}'')$ be its value in directions $\hat{Z}',\hat{Z}''$.[11] It would be helpful to remember that the second-order directional

[9] In particular, in our example of the neoclassical growth model, $\mathsf{P} = [1, 0, 0]$.

[10] Given the simple way in which $\{\hat{Y}_t\}_{t=0}^T$ depends on $\{\hat{X}_t\}_{t=0}^T$, this inversion is particular easy to do.

[11] To build intuition for these objects, consider our neoclassical growth model example. We have $\overline{K}:\mathbb{R}^2\to\mathbb{R}$, and so $\overline{K}_{ZZ}$ is a 2×2 matrix of cross-partial derivatives of $\overline{K}$ with respect to Θ and K. Directional derivative $\overline{K}_{ZZ}\cdot(\hat{Z}',\hat{Z}'')$ is a scalar computed as $(\hat{Z}'')^{\mathrm{T}}\overline{K}_{ZZ}\hat{Z}'$. It captures the second order interaction effect of changes $\hat{Z}'$ and $\hat{Z}''$ on capital $\overline{K}$. Vector $\overline{X}$ consists of three policy functions, so $\overline{X}:\mathbb{R}^2\to\mathbb{R}^3$, $\overline{X}_{ZZ}$ is a collection of three 2×2 matrices of cross-partial derivatives and $\overline{X}_{ZZ}\cdot(\hat{Z}',\hat{Z}'')$ is a 3×1 vector that captures the interaction effect of $\hat{Z}'$ and $\hat{Z}''$ on each of the three policy functions.

derivative of compounded functions satisfies the following relationship

$$\begin{aligned}&\lim_{\alpha''\to 0}\lim_{\alpha'\to 0}\frac{1}{\alpha''}\frac{1}{\alpha'}\left[\overline{X}\left(\overline{Z}\left(Z^*+\alpha'\hat{Z}'+\alpha''\hat{Z}''\right)\right)-\overline{X}\left(\overline{Z}\left(Z^*\right)\right)\right] \\ =&\overline{X}_Z\cdot\overline{Z}_{ZZ}\cdot(\hat{Z}',\hat{Z}'')+\overline{X}_{ZZ}\cdot(\overline{Z}_Z\cdot\hat{Z}',\overline{Z}_Z\cdot\hat{Z}'').\end{aligned} \tag{11}$$

Using this observation, we take the second-order Taylor expansion of (8) and obtain the second-order analogue of Lemma 1:

Lemma 2. *To the second-order approximation, X_t satisfies*

$$X_t\left(\mathcal{E}^t\right)=...+\frac{1}{2}\left(\sum_{s=0}^{t}\sum_{m=0}^{t}\hat{X}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m+\hat{X}_{\sigma\sigma,t}\right)+O\left(\|\mathcal{E}\|^3\right), \tag{12}$$

where ... are the first-order terms and $\left\{\hat{X}_{t,k}\right\}_{t,k}$ and $\left\{\hat{X}_{\sigma\sigma,t}\right\}_t$ are defined by

$$\hat{X}_{t,k}:=\overline{X}_Z\cdot\hat{Z}_{t,k}+\overline{X}_{ZZ}\cdot\left(\hat{Z}_t,\hat{Z}_k\right)\ \textit{for}\ \hat{Z}_{t,k}=\overline{Z}_Z\cdot\hat{Z}_{t-1,k-1}+\overline{Z}_{ZZ}\cdot\left(\hat{Z}_{t-1},\hat{Z}_{k-1}\right), \tag{13}$$

$$\overline{X}_{\sigma\sigma,t}:=\overline{X}_Z\cdot\hat{Z}_{\sigma\sigma,t}+\overline{X}_{\sigma\sigma}\ \textit{for}\ \hat{Z}_{\sigma\sigma,t}=\overline{Z}_Z\cdot\hat{Z}_{\sigma\sigma,t-1}+\left[0,\mathrm{P}\overline{X}_{\sigma\sigma}\right]^{\mathrm{T}}. \tag{14}$$

with $\hat{Z}_{0,k}=\hat{Z}_{t,0}=\hat{Z}_{\sigma\sigma,0}=0$.

This lemma shows that the second-order equilibrium approximation involves two types of terms: terms like $\hat{X}_{t,k}$ that capture the interaction effects on current period endogenous variables from shocks that occurred t and k period ago, and terms like $\hat{X}_{\sigma\sigma,t}$ that capture precautionary motives. Inspection of equation (13) reveals that both $\hat{Z}_{t,k}$ and $\hat{X}_{t,k}$ have the same mathematical structure as our example in equation (11). Equation (14) is simpler since precautionary motives are zero to the first-order, so that the first-order interaction terms drop out.

To find $\{\hat{X}_{t,k}\}_{t,k}$ and $\{\hat{X}_{\sigma\sigma,t}\}_t$, we proceed the same way as we did in the first-order approximations and differentiate (5) twice with respect to σ. Let G_{YY} be the Hessian of G evaluated at $Y=Y^*$. A helpful observation to build intuition for the next proposition is that

$$\left.\frac{\partial^2}{\partial\sigma^2}\mathbb{E}_{\mathcal{E}}\overline{X}\left(\sigma\mathcal{E},K^*;\sigma\right)\right|_{\sigma=0}=\mathbb{E}_{\mathcal{E}}\left[\overline{X}_{ZZ}\cdot\left([\mathcal{E},0]^{\mathrm{T}},[\mathcal{E},0]^{\mathrm{T}}\right)\right]+\overline{X}_{\sigma\sigma}=\hat{X}_{0,0}var\left(\mathcal{E}\right)+\overline{X}_{\sigma\sigma}.$$

Proposition 2. *(a). For any k, sequence $\left\{\hat{X}_{t,t+k}\right\}_t$ satisfies*

$$\mathsf{G}_Y\hat{Y}_{t,t+k}+\hat{\mathsf{G}}_{t,t+k}=0\ \textit{for all}\ t, \tag{15}$$

where $\hat{\mathsf{G}}_{t,t+k}=\mathsf{G}_{YY}\cdot\left(\hat{Y}_t,\hat{Y}_{t+k}\right)$, $\hat{Y}_{t,t+k}=[0,\mathsf{P}\hat{X}_{t-1,t+k-1},\hat{X}_{t,t+k},\hat{X}_{t+1,t+k+1}]^{\mathrm{T}}$, $\mathsf{P}\hat{X}_{-1,k-1}=0$, and $\lim_{t\to\infty}\hat{X}_{t,t+k}=0$.

(b). Sequence $\left\{\hat{X}_{\sigma\sigma,t}\right\}_t$ satisfies

$$\mathsf{G}_Y\hat{Y}_{\sigma\sigma,t}=0\ \textit{for all}\ t, \tag{16}$$

where $\hat{Y}_{\sigma\sigma,t}=\left[0,\mathsf{P}\hat{X}_{\sigma\sigma,t-1},\hat{X}_{\sigma\sigma,t},\hat{X}_{\sigma\sigma,t+1}+\hat{X}_{0,0}var\left(\mathcal{E}\right)\right]^{\mathrm{T}}$, $\mathsf{P}\hat{X}_{\sigma\sigma,-1}=0$ and $\lim_{t\to\infty}\hat{X}_{\sigma\sigma,t}-\hat{X}_{\sigma\sigma,t-1}=0$.

Similar to Proposition 1, Proposition 2 collapses the problem of finding second-order approximation to solving systems of linear equations. To solve $\{\hat{X}_{t,t+k}\}_t$ for any given k we use (15). It has mathematical structure similar to (10) except it requires contracting $\hat{\mathsf{G}}_{t,t+k}$. To construct it, we first find G_{YY} by automatically differentiating (2) twice and evaluating it at steady state $Y = Y^*$, and then using the solution $\{\hat{Y}_t\}_t$ obtained in Proposition 1 to compute $\{\hat{\mathsf{G}}_{t,t+k}\}_t$. In the same manner as the first order, the stability assumption 1b can be used to show the boundary condition $\lim_{t\to\infty} \hat{X}_{t,t+k} = 0$. The systems of equations (15) are therefore solved using truncation and a terminal condition $\hat{X}_{t,t+k} = 0$ for $t > T$. These systems of equations can be easily parallelized and solved simultaneously for all k. Once we have $\hat{X}_{0,0}$, we can solve solve (16) for $\{\hat{X}_{\sigma\sigma,t}\}_t$. The only difference is that this system of equations has the boundary condition $\lim_{t\to\infty} \hat{X}_{\sigma\sigma,t} - \hat{X}_{\sigma\sigma,t-1} = 0$ which we implement with the terminal condition $\hat{X}_{\sigma\sigma,t} = \hat{X}_{\sigma\sigma,T}$ for $t > T$, since precautionary motives do not need to die off as $t \to \infty$.

As with the first-order approximations, our implementation of perturbational techniques avoid common complications that arise under standard techniques. We do not need to compute and store matrices of second-order partial derivatives such as $\overline{X}_{ZZ}$ or to implement additional "pruning" techniques to find stable solutions.[12]

Lemma 2 can be used to find the approximation to the ergodic mean of endogenous variables in the stochastic economy. Taking expectation of equation (12) and then the limit as $t \to \infty$ finds the long run average level of X to be

$$\mathbb{E}[X] = \overline{X} + \sum_{s=0}^{\infty} \hat{X}_{s,s} var(\mathcal{E}) + \lim_{t\to\infty} \hat{X}_{\sigma\sigma,t} + O\left(\|\mathcal{E}\|^3\right). \tag{17}$$

If aggregate welfare is included in X then equation (17) can be used to compute ergodic welfare. The observation that the difference $\mathbb{E}[X] - \overline{X}$ is of the second order implies that the approximation error would not improve if one were to approximate the equilibrium dynamics around the ergodic mean of K rather than deterministic steady-state K^*.

3 Approximations of HA economies

The directional derivative approach is particularly convenient to approximate HA models with aggregate shocks. In those models, state Z depends on the distribution of individual characteristics, such as asset holdings, and usually is infinite dimensional. On the other hand, one is typically interested in finding a small number of endogenous variables, such as aggregate prices and quantities or some moments summarizing heterogeneity. This makes the object of interest X small dimensional. The techniques

[12] It is well-known that standard perturbational methods produce unstable sample paths at higher orders. Kim et al. (2008) propose to apply a "pruning" procedure to the obtained solution to address that issue. Lombardo and Uhlig (2018) use series expansion methods of Holmes (2012) and Lombardo (2010) to provide a systematic theory of pruning. Our second-order solution coincides with the one described in Lombardo and Uhlig (2018).

that we developed in the previous section allow us to compute the responses of X quickly for arbitrary Z. To make our discussion concrete, we start by describing how we would approach solving a canonical HA economy, the one considered by Krusell and Smith (1998).

3.1 Prototypical Krusell and Smith economy

The economy is populated by a continuum of households that face idiosyncratic risk and firms. Each household supplies inelastically one unit of labor that is subject to idiosyncratic efficiency shocks $\theta_{i,t}$. Households receive wage W_t and save capital $k_{i,t}$ that earns gross return R_t, that equals rental rate net of depreciation δ. Household i chooses stochastic sequences $\{c_{i,t}, k_{i,t}\}_t$ to maximize life-time expected utility $\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U(c_{i,t})$ subject to the budget constraints $c_{i,t} + k_{i,t} \leq R_t k_{i,t-1} + W_t \exp(\theta_{i,t})$ and the borrowing constraints $k_{i,t} \geq 0$ for all $t \geq 0$. Initial $k_{i,-1}$ and $\theta_{i,0}$ are given and the distribution of $\{k_{i,-1}, \theta_{i,0}\}_i$ over i is denoted by Ω_0. Efficiency $\theta_{i,t}$ follows an exogenous stationary stochastic process normalized so that $\int \exp(\theta_{i,t}) di = 1$.

Households rent capital and supply efficiency-adjusted labor to firms each period. Firms are competitive and produce output using Cobb-Douglas technology with aggregate productivity $\exp(\Theta_t)$ and a capital share of α. Wages W_t and rental rates R_t are determined by the market clearing conditions so that supply of labor and capital by consumers is equal to the demand for those factors by firms.

The equilibrium in this economy can be represented by three set of conditions: the optimality conditions of households that face idiosyncratic risk, the optimality conditions of firms, and the market clearing conditions. It is helpful to keep conditions characterizing behavior of economic agents that face idiosyncratic risk (i.e., households) separate from the other conditions. Let $\zeta_{i,t}$ be the Lagrange multiplier on the borrowing constraint of household i in period t. We can write the optimality conditions of households as

$$R_t U_c(c_{i,t}) - \lambda_{i,t} = 0, \quad U_c(c_{i,t}) + \zeta_{i,t} - \beta \mathbb{E}_t \lambda_{i,t+1} = 0 \text{ for all } i, t, \tag{18}$$

and

$$c_{i,t} + k_{i,t} - R_t k_{i,t-1} - W_t \exp(\theta_{i,t}) = 0, \quad k_{i,t} \zeta_{i,t} = 0 \text{ for all } i, t. \tag{19}$$

Here, equations (18) are households' Euler equations and equations (19) are budget constraints and the complementary slackness conditions on borrowing constraints. Besides equations (18) and (19), we have non-negativity constraints $k_{i,t} \geq 0$, $\zeta_{i,t} \geq 0$ for all i, t.

Letting A_t represent capital supplied by the households in date t, the optimality conditions of firms and market clearing conditions are

$$W_t - (1-\alpha) \exp(\Theta_t) K_t^{\alpha} = 0, \quad K_t - A_{t-1} = 0 \text{ for all } t, \tag{20}$$

$$R_t + \delta - (1-\alpha)\exp(\Theta_t) K_t^{\alpha-1} = 0, \quad A_t - \int k_{i,t} di = 0 \text{ for all } t. \tag{21}$$

Given initial conditions $\{k_{i,-1}, \theta_{i,0}\}_i$ and Θ_0, equations (18), (19), (20) and (21), and non-negativity constraints characterize equilibrium dynamics of this economy.

3.2 General representation of HA economies

Motivated by this example, we now present a general representation of equilibrium conditions of a broad class of HA economies. Let θ and Θ be vectors of idiosyncratic and aggregate shocks. Let x be the vector of endogenous variables chosen by the agents subject to idiosyncratic shocks and X be the vector of all other endogenous variables. We refer to x and X as idiosyncratic and aggregate endogenous variables. Let $a_{i,t-1} \in x_{i,t-1}$ and $A_{t-1} \in X_{t-1}$ be vectors of individual and aggregate endogenous variables chosen at $t-1$ that enter into time t equilibrium conditions. We will write them explicitly as

$$a = \mathsf{p}x, \quad A = \mathsf{P}X,$$

where p and P are selection matrices that return a and A from vectors x and X. In what follows, we will use A and $\mathsf{P}X$ interchangeably. Let $Y_t := [\Theta_t, \mathsf{P}X_{t-1}, X_t, \mathbb{E}_t X_{t+1}]^{\mathrm{T}}$.

The optimality conditions of agents subject to idiosyncratic risk are represented as

$$F(a_{i,t-1}, \theta_{i,t}, x_{i,t}, \mathbb{E}_{i,t} x_{i,t+1}, Y_t) = 0 \text{ for all } i, t \tag{22}$$

with initial conditions $(a_{i,-1}, \theta_{i,0})$. Let Ω_0 the the (cumulative) distribution of $(a_{i,-1}, \theta_{i,0})$. The remaining equilibrium conditions, which include optimality conditions of agents not subject to idiosyncratic shocks, market clearing conditions, budget constraints for the government, etc, are represented as

$$G\left(Y_t, \int x_i di\right) = 0 \text{ for all } t \tag{23}$$

with some initial Θ_0 and A_{-1}.

It is easy to see how our example of the Krusell and Smith economy fits into this representation. In that example, we have $x_{i,t} = [k_{i,t}, c_{i,t}, \lambda_{i,t}, \zeta_{i,t}]^{\mathrm{T}}$, $a_{i,t} = k_{i,t}$, $A_t = \int a_{i,t} di$, $X_t = [A_t, K_t, W_t, R_t]^{\mathrm{T}}$, $Y_t = [\Theta_t, A_{t-1}, X_t, \mathbb{E}_t X_{t+1}]$. With these definitions, mapping F captures optimality conditions (18) and (19), while mapping G captures conditions (20) and (21).

In order to streamline our exposition, for now we assume that $\theta_{i,t}$, Θ_t and $a_{i,t}$ are scalars; we discuss the case when they are finite vectors in the supplementary material Section B. We assume that $\theta_{i,t}$ follows AR(1) processes

$$\theta_{i,t} = \rho_\theta \theta_{i,t-1} + \varepsilon_{i,t}, \tag{24}$$

and Θ_t is given by (1). Here, $\varepsilon_{i,t}$ is a mean zero random variable drawn independently across time and agents with a probability distribution that has a density μ and $|\rho_\theta| < 1$.

An *equilibrium* consists of stochastic processes $\{X_t(\mathcal{E}^t)\}_{t,\mathcal{E}^t}$ and $\{x_{i,t}(\mathcal{E}^t, \varepsilon_i^t)\}_{i,t,\mathcal{E}^t,\varepsilon^t}$ that satisfy (22)–(24) and auxiliary non-negativity constraints given initial conditions $Z_0 = [\Theta_0, A_{-1}, \Omega_0]$. Our main focus is on characterizing the equilibrium stochastic process $\{X_t\}_t$, which is relevant for most macroeconomic applications. As a by-product, we also describe a procedure to recover the stochastic processes $\{x_{i,t}\}_{i,t}$.

We proceed as in Section 2 and perturb the aggregate shock process as in (4). In a recursive formulation, the aggregate state of the system (22) and (23) consists of Θ_t, A_{t-1} and the joint distribution Ω_t over $\{(a_{i,t-1}, \theta_{i,t})\}_i$. We use $\Omega_t\langle a, \theta\rangle$ to denote the measure of agents with $\theta_{i,t} \leq \theta$ and $a_{i,t-1} \leq a$. Let $Z_t = [\Theta_t, A_{t-1}, \Omega_t]^{\mathrm{T}}$ be the aggregate state. The recursive representation of equilibrium conditions in the perturbed economy is given by

$$F\left(a, \theta, \overline{x}(a, \theta, Z; \sigma), \mathbb{E}_{\varepsilon,\mathcal{E}}[\overline{x}|a, \theta, Z; \sigma], \overline{Y}(Z; \sigma)\right) = 0 \text{ for all } (a, \theta, Z, \sigma), \tag{25}$$

$$G\left(\overline{Y}(Z; \sigma), \int \overline{x}(\cdot, \cdot, Z; \sigma)\, d\Omega\right) = 0 \text{ for all } Z, \sigma, \tag{26}$$

as well as the LoM for the aggregate distribution $\overline{\Omega}(Z; \sigma)$ defined as

$$\overline{\Omega}(Z; \sigma)\langle a', \theta'\rangle = \int\int \iota\left(\overline{a}(a, \theta, Z; \sigma) \leq a'\right) \iota(\rho_\theta \theta + \varepsilon \leq \theta')\mu(\varepsilon)\, d\varepsilon d\Omega\langle a, \theta\rangle \text{ for all } Z, \sigma, \tag{27}$$

where $\overline{Y}(Z; \sigma) = \left[\Theta, A, \overline{X}(Z; \sigma), \mathbb{E}_{\mathcal{E}}\left[\overline{X}|Z; \sigma\right]\right]^{\mathrm{T}}$ and $\mathbb{E}_{\varepsilon,\mathcal{E}}$ and $\mathbb{E}_{\mathcal{E}}$ to denote conditional expectation of future policies with respect to $(\varepsilon, \mathcal{E})$ and $\mathcal{E}$, respectively:

$$\mathbb{E}_{\varepsilon,\mathcal{E}}[\overline{x}|a, \theta, Z; \sigma] = \int \overline{x}\left(\overline{a}(a, \theta, Z; \sigma), \rho_\theta\theta + \varepsilon, \rho_\Theta\Theta + \sigma\mathcal{E}, \mathsf{P}\overline{X}(Z; \sigma), \overline{\Omega}(Z; \sigma)\right) \mu(\varepsilon)\, d\varepsilon d\Pr(\mathcal{E}),$$

$$\mathbb{E}_{\mathcal{E}}\left[\overline{X}|Z; \sigma\right] = \int \overline{X}\left(\rho_\Theta\Theta + \sigma\mathcal{E}, \mathsf{P}\overline{X}(Z; \sigma), \overline{\Omega}(Z; \sigma)\right) d\Pr(\mathcal{E}).$$

This recursive system is initialized by the initial condition Z_0.[13]

We proceed as in Section 2 and approximate equilibrium responses around the steady state of the deterministic economy.[14] We denote this steady state by $Z^* = [0, A^*, \Omega^*]^{\mathrm{T}}$. Let $\overline{\Lambda}(a', \theta', a, \theta)$ be the transition kernel from (a, θ) to (a', θ') in the steady state. As in Section 2, we drop explicit dependence of policy functions on σ and Z when $\sigma = 0$ and $Z = Z^*$. Thus, for example, $\overline{x}(a, \theta)$ will be understood as $\overline{x}(a, \theta, Z^*; 0)$. $\overline{Z}(Z) := [\rho_\Theta\Theta, \overline{A}(Z), \overline{\Omega}(Z)]^{\mathrm{T}}$ is the LoM for the aggregate state in the deterministic economy. As before, $\overline{X}_Z$, $\overline{X}_{ZZ}$, $\overline{Z}_Z$, $\overline{x}_Z(a, \theta)$, etc, denotes Fréchet derivatives of policy functions of various orders. In HA settings, these are infinite dimensional linear operators rather than finite

[13]We do not explicitly add the non-negativity constraints in our state-space formulation. As we show formally in the appendix, those constraints are not used in our perturbations. The reason for why they are not needed is the following. We start with a steady state economy that already satisfies those inequalities and then perturb it. All strict inequalities in the steady state remain strict in the perturbed economy and thus do not affect our analysis. The inequalities that hold as equalities are already incorporated explicitly in F mapping as complementary slackness conditions, and our approach automatically calculates how perturbations affect them.

[14]As in Section 2, we refer to the economy with $\sigma = 0$ as deterministic economy. Note that agents are still subject to idiosyncratic shocks in this economy. Ω^* is the invariant distribution in such economy.

dimensional matrices that we used in Section 2, but their analytical properties used in the proofs are largely unaffected by this distinction.

Our approach to approximating the HA economy will be parallel to the way we the approximated RA economy in Section 2. We assume that aggregate policy functions $\overline{X}$ are sufficiently smooth and stable, in the sense of Assumption 1, and use various orders of Taylor expansions of (25)–(27) to find first- and second-order approximations. At the same time, many HA economies have features that are not present in the standard applications of perturbational techniques, namely kinks in policy functions $\overline{x}$ arising, for example, due to the occasional binding constraints. To allow for such kinks, we impose weaker conditions on $\overline{x}$.

Assumption 2.

(a) $\overline{x}\,(a,\theta,Z;\sigma)$ is continuous and piecewise sufficiently differentiable at $(Z^*,0)$ for all (a,θ). The points of non-differentiability of $\overline{x}\,(\cdot,\theta,Z;\sigma)$ are described by a finite number of sufficiently differentiable functions $\{\overline{\kappa}_j(\theta,Z;\sigma)\}_j$;

(b) The marginal distribution $\int \Omega^* d\theta$ has a finite number of mass-points $\{a_n^*\}_n$, i.e., Ω^* has compact support and has density $\mathring{\omega}^*(a,\theta)+\sum_n \xi_n^*(\theta)\delta(a-a_n^*)$, where $\mathring{\omega}^*(a,\theta)$ and $\xi_n^*(\theta)$ are continuous and δ is a Dirac delta function.

Condition (a) allows policy functions to have a finite number of kinks, with j^{th} kink of $\overline{x}(\cdot,\theta,Z;\sigma)$ occurring at $a=\overline{\kappa}_j(\theta,Z;\sigma)$. Note that the kinks are endogenous in the sense that they depend on the aggregate state Z. Condition (b) permits the marginal of the invariant distribution Ω^* to have a finite number of mass points. A direct implication of conditions (a) and (b) is that the set of points for which $\overline{x}$ is not differentiable at $(Z^*,0)$ is of Ω^*-measure zero. This also implies that the integral $\int \overline{x}d\Omega$ is differentiable at $(Z^*,0)$.

As an illustration, consider our example of the Krusell and Smith economy. In that economy, policy function $\overline{k}(k,\theta)$ is continuous, strictly increasing for $\theta \geq \overline{\theta}^{\vee}(k)$ and equal to zero for $\theta \leq \overline{\theta}^{\vee}(k)$, where $\overline{\theta}^{\vee}(k)$ is the level of θ at which the borrowing constraint starts to bind. Since the distribution of idiosyncratic shocks has a density μ, the marginal of Ω^* can have at most one mass point, at $k=0$. Thus, its density takes the form $\mathring{\omega}^*(k,\theta)+\xi^*(\theta)\delta(k)$, where $\xi^*\,(\theta)$ is the density of agents with productivity θ at $k=0$. Function $\overline{\kappa}\,(\theta)$ is the inverse of $\overline{\theta}^{\vee}(k)$.

For the rest of the paper, we assume that policy function $\overline{x}\,(\cdot,\theta)$ has at most one kink, at some $\overline{\kappa}(\theta)$. This done merely to simplify the exposition. The extension of our formulas to accommodate finite number of kinks is immediate.

We use $\overline{x}_a(a,\theta)$ and $\overline{x}_{aa}(a,\theta)$ to represent derivatives of policy functions with respect to a. Since policy functions may have kinks, these derivatives are not defined in the classical sense at those kinks. To handle this, we treat all derivatives of individual policy functions as generalized or distributional derivatives and represent them as generalized functions.[15] Since generalized functions are infinitely differentiable (in the distributional derivative sense), they allow us to present a uniform treatment of approximations at any order. When we want to distinguish between generalized and classical functions, we use symbol $^\circ$ to denote the latter. To see the relationship between the two, let $\overline{x}^{\Delta}(\theta) = \lim_{a\downarrow\overline{\kappa}(\theta)} \overline{x}(a,\theta) - \lim_{a\uparrow\overline{\kappa}(\theta)} \overline{x}(a,\theta)$ and let $\overline{x}_a^{\Delta}(\theta)$ be defined analogously for $\overline{x}_a(\cdot,\theta)$. Obviously, $\overline{x}^{\Delta}(\theta) = 0$ by continuity of $\overline{x}(\cdot,\theta)$ but $\overline{x}_a^{\Delta}(\theta) \neq 0$ due to kinks. The relationship between generalized and classical derivatives is given by

$$\overline{x}_a(a,\theta) = \mathring{\overline{x}}_a(a,\theta) + \underbrace{\delta\left(a - \overline{\kappa}(\theta)\right)\overline{x}^{\Delta}(\theta)}_{=0},$$
$$\overline{x}_{aa}(a,\theta) = \mathring{\overline{x}}_{aa}(a,\theta) + \delta\left(a - \overline{\kappa}(\theta)\right)\overline{x}_a^{\Delta}(\theta).$$

These relationships imply that integrals of $\overline{x}_a$ and $\mathring{\overline{x}}_a$ always agree, but integrals of $\overline{x}_{aa}$ differ from $\mathring{\overline{x}}_{aa}$ by terms involving jumps at the kinks, e.g.,

$$\int \overline{x}_{aa} d\Omega^* = \int \mathring{\overline{x}}_{aa} d\Omega^* + \int \overline{x}_a^{\Delta}(\theta)\,\omega^*\left(\overline{\kappa}(\theta),\theta\right) d\theta,$$

where $\omega^*(a,\theta) := \mathring{\omega}^*(a,\theta) + \sum_n \xi_n^*(\theta)\delta(a - a_n^*)$. Note that ω^* is a generalized function as well and notations such as $\int \overline{x} d\Omega^*$ and $\int \overline{x}\omega^* d\theta da$ are equivalent under this convention.

Derivatives of individual policy functions such as $\overline{x}_Z(a,\theta)\cdot\hat{Z}$ or $\overline{x}_{ZZ}(a,\theta)\cdot(\hat{Z}',\hat{Z}'')$ and cross-partials $\overline{x}_{aZ}(a,\theta)\cdot\hat{Z}$ also will be understood to be represented by generalized functions. We show in the appendix that they satisfy $\overline{x}_Z(a,\theta)\cdot\hat{Z} = \mathring{\overline{x}}_Z(a,\theta)\cdot\hat{Z}$ and

$$\overline{x}_{aZ}(a,\theta)\cdot\hat{Z} = \mathring{\overline{x}}_{aZ}(a,\theta)\cdot\hat{Z} + \delta(a-\overline{\kappa}(\theta))\overline{x}_Z^{\Delta}(\theta)\cdot\hat{Z}, \tag{28}$$
$$\overline{x}_{ZZ}(a,\theta)\cdot\left(\hat{Z}',\hat{Z}''\right) = \mathring{\overline{x}}_{ZZ}(a,\theta)\cdot\left(\hat{Z}',\hat{Z}''\right) + \delta(a-\overline{\kappa}(\theta))\overline{x}_a^{\Delta}(\theta)\left(\overline{\kappa}_Z(\theta)\cdot\hat{Z}'\right)\left(\overline{\kappa}_Z(\theta)\cdot\hat{Z}''\right).$$

Term $\overline{\kappa}_Z(\theta)\cdot\hat{Z}$ in the second equation represents how the kink moves when the aggregate state is changed in direction $\hat{Z}$.

3.3 First-order approximations

We start with the first-order approximations. For the HA economy, the statement (and the proof) of Lemma 1 remains unchanged as long as we augment the LoM for $Z_t(\mathcal{E}^t;\sigma)$ in equation (9) to include $\overline{\Omega}\left(Z_{t-1}\left(\mathcal{E}^{t-1};\sigma\right);\sigma\right)$ and understand $\hat{Z}_0 = [1,0]^{\mathrm{T}}$ to mean the infinite-dimensional vector that consists of 1 in the first element (corresponding to the exogenous part of the state space) and 0 for all other elements. To characterize $\{\hat{X}_t\}_t$ we differentiate (26) and evaluate it in direction $\hat{Z}_t = \left[\rho_\Theta^t, \mathsf{P}\hat{X}_{t-1}, \hat{\Omega}_t\right]^{\mathrm{T}}$

[15] A generalized function is a linear functional over some space of functions. For instance, δ is a generalized function defined by the operation $\delta[\phi] = \int \phi(x)\delta(x)dx = \phi(0)$ for some function ϕ. There is a large mathematical literature on generalized functions (also referred to as distributions) and distributional derivatives, see Kanwal (1998) for an introduction to this subject.

to obtain

$$\mathsf{G}_Y \hat{Y}_t + \mathsf{G}_x \left(\int \overline{x} d\Omega \right)_Z \cdot \hat{Z}_t = 0, \tag{29}$$

where $\hat{Y}_t$ is as defined in Proposition 1. Here, G_x is the derivative of G with respect to its second argument, $\int x d\Omega$, evaluated at the steady state, and $\left(\int \overline{x} d\Omega\right)_Z \cdot \hat{Z}_t$ is the directional derivative of $\int \overline{x} d\Omega$ in direction $\hat{Z}_t$. Using the definition of directional derivatives and continuity of policy functions at the kinks, it is easy to show that

$$\left(\int \overline{x} d\Omega \right)_Z \cdot \hat{Z}_t = \int \hat{x}_t d\Omega^* + \int \overline{x} d\hat{\Omega}_t, \tag{30}$$

where $\hat{x}_t(a,\theta) = \overline{x}_Z(a,\theta) \cdot \hat{Z}_t$ are directional derivatives of individual policy functions.

Equation (30) shows that the first-order change in the aggregation $\int \overline{x} d\Omega$ consists of two components: the effect of the shock on individual policy functions, $\int \hat{x}_t d\Omega^*$, and the effect of the shock on the aggregate distribution, $\int \overline{x} d\hat{\Omega}_t$. In order to characterize these components, we use two intermediate results. First, we show that there is a tight relationship between responses of individual and aggregate endogenous variables to aggregate shocks. We obtain it by applying the implicit function theorem to mapping F defined in equation (25) and evaluating those expressions at the deterministic steady-state.

Lemma 3. *For any t,*

$$\hat{x}_t(a,\theta) = \sum_{s=0}^{\infty} \mathsf{x}_s(a,\theta) \hat{Y}_{t+s}, \tag{31}$$

where matrices $\mathsf{x}_s(a,\theta)$ are given by

$$\mathsf{x}_0(a,\theta) = -\left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon[\overline{x}_a|a,\theta]\,\mathsf{p}\right)^{-1} \mathsf{F}_Y(a,\theta), \tag{32}$$

$$\mathsf{x}_{s+1}(a,\theta) = -\left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon[\overline{x}_a|a,\theta]\,\mathsf{p}\right)^{-1} \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon[\mathsf{x}_s|a,\theta] \tag{33}$$

away from the kinks and $\mathsf{x}_s(\overline{\kappa}(\theta),\theta) = 0$ at the kinks, and $\mathsf{F}_x(a,\theta)$, $\mathsf{F}_{x^e}(a,\theta)$, $\mathsf{F}_Y(a,\theta)$ are derivatives of F with respect to x, $\mathbb{E}x$ and Y, all evaluated at the steady state values of $\overline{x}(a,\theta)$.

Equation (31) shows that the change in individual policy functions $\hat{x}_t$ is equal to the future changes in aggregates $\{\hat{Y}_{t+s}\}_s$ weighted with matrices $\{\mathsf{x}_s\}_s$. Matrix x_s has a natural economic interpretation. It captures how much individuals change their policy functions today if they expect aggregates to change s periods in the future, $\partial x_t / \partial Y_{t+s}$. The most important part of Lemma 3 is that it provides explicit formulas for $\{\mathsf{x}_s\}_s$. We show in Section 4.1 that, similarly to finding G_Y in Section 2, matrices F_x, F_{x^e}, F_Y are easy to obtain by automatically differentiating mapping F and evaluating it at the steady state. Thus, the right hand side of (32) is known from zeroth-order terms, and therefore, x_0 can be constructed using linear algebra operations. This allows the construction of $\{\mathsf{x}_s\}_{s>0}$ sequentially using (33).

The second intermediate result describes the Law of Motion for $\hat{\Omega}_t$, that helps simplifying the second term on the right hand side of (30). It is helpful to define three operators, $\mathcal{M}$, $\mathcal{L}^{(a)}$, and $\mathcal{I}^{(a)}$ that return, for any generalized function y, the following objects:

$$\begin{aligned}(\mathcal{M}\cdot y)\langle a',\theta'\rangle &:= \int \overline{\Lambda}(a',\theta',a,\theta)\omega^*(a,\theta)\,y(a,\theta)\,dad\theta,\\ \left(\mathcal{L}^{(a)}\cdot y\right)\langle a',\theta'\rangle &:= \int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\,y(a,\theta)\,dad\theta,\\ \mathcal{I}^{(a)}\cdot y &:= \int \overline{x}_a(\theta,a)y(\theta,a)\,dad\theta.\end{aligned}$$

Operators $\mathcal{M}$ and $\mathcal{L}^{(a)}$ take a function y, multiply it by ω^* and $\overline{a}_a$ respectively, and integrate that product over $\overline{\Lambda}(a',\theta',\cdot,\cdot)$ for a given value of (a',θ'). Operator $\mathcal{I}^{(a)}$ is an integral of the product $\overline{x}_a y$. To explain the economic forces captured by these operators, we start with the following lemma, which we obtain by explicitly taking derivatives of the LoM, $\overline{\Omega}$, defined in equation (27).

Lemma 4^{HA}. *For any t, $\frac{d}{d\theta}\hat{\Omega}_t$ satisfies a recursion with $\frac{d}{d\theta}\hat{\Omega}_0=\mathbf{0}$ and*

$$\frac{d}{d\theta}\hat{\Omega}_{t+1}=\mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_t-\mathcal{M}\cdot\hat{a}_t, \tag{34}$$

Equation (34) describes how the aggregate distribution Ω_t is affected by aggregate shocks. On impact of the shock in period 0, the distribution is pre-determined and thus $\frac{d}{d\theta}\hat{\Omega}_0=\mathbf{0}$. Individuals change their choices in period 0. In particular, individual (a,θ) changes her savings behavior by $\hat{a}_0(a,\theta)$. Operator $\mathcal{M}$ aggregates these individual-level changes by weighting them with the invariant density ω^* and returns the change in the distribution in period 1, $\frac{d}{d\theta}\hat{\Omega}_1=-\mathcal{M}\cdot\hat{a}_0$. Thus, $\mathcal{M}$ captures the first-order effect of changes in individual policy functions on the aggregate distribution next period. For all $t>0$, the distribution Ω_t is affected by two forces. One is mechanical: if the distribution Ω_{t-1} changed in the previous period, Ω_t would also change even if individual policy functions did not change. This mechanical effect is captured by the operator $\mathcal{L}^{(a)}$. The aggregate distribution in period $t+1$ is also affected by the response of individuals in period t, and this behavioral effect is captured by $\mathcal{M}\cdot\hat{a}_t$.

To make recursion (34) operational, note that $\int\overline{x}d\hat{\Omega}_t=-\mathcal{I}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_t$ using integration by parts. This leads to the following corollary that characterizes the derivative of the aggregation equation (30).

Corollary 1. *For any t,*

$$\left(\int\overline{x}d\Omega\right)_Z\cdot\hat{Z}_t=\sum_{s=0}^{\infty}\mathsf{J}_{t,s}\hat{Y}_s$$

where $\{\mathsf{J}_{t,s}\}_{t,s}$ satisfies a recursion with $\mathsf{J}_{0,s}=\int\mathsf{x}_s d\Omega^*$ and

$$\mathsf{J}_{t,s}=\mathsf{J}_{t-1,s-1}+\mathcal{I}^{(a)}\cdot\left(\mathcal{L}^{(a)}\right)^{t-1}\cdot\mathcal{M}\cdot\mathsf{px}_s, \tag{35}$$

with the convention $\mathsf{x}_k=0$ for $k<0$.

This corollary shows that relevant directional derivatives of $\int \overline{x} d\Omega$ can be expressed purely in terms of $\{\hat{Y}_t\}_t$ weighted with matrices $\{\mathsf{J}_{t,s}\}_{t,s}$ that is described by a linear recursive system of equations (35). Combine Corollary 1 and equation (29) to obtain the HA analogue of our Proposition 1.

Proposition 1^{HA}. $\{\hat{X}_t\}_t$ *satisfies*

$$\mathsf{G}_Y \hat{Y}_t + \mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \hat{Y}_s = 0 \textit{ for all } t, \tag{36}$$

where $\hat{Y}_t$ *is as defined in Proposition 1,* $\mathsf{P}\hat{X}_{-1} = 0$, *and* $\lim_{t\to\infty} \hat{X}_t = 0$.

Relative to the economy considered in Section 2, heterogeneity adds an additional term, captured by the infinite sum in (36). In order to compute it, we need to compute $\{\mathsf{x}_t\}_t$ using (32) and (33), construct operators $\mathcal{M}$, $\mathcal{L}^{(a)}$, and $\mathcal{I}^{(a)}$, and then compute $\{\mathsf{J}_{t,s}\}_{t,s}$ using (35). As we show in Section 4, all these steps can be done very easily using the output of the standard off-the-shelf algorithms that compute steady states of the HA economies without aggregate shocks. Once this step is completed, one can find $\{\hat{X}_t\}_t$ in the same way we did it in Section 2.

After $\{\hat{X}_t\}_t$ is constructed, other objects can be easily recovered. In particularly, $\{\hat{x}_t\}_t$ can be computed using (31). Similarly, one can show (see Appendix B.2) that continuity of policy functions implies the effect of aggregate shocks on the policy function that describes kinks, $\overline{\kappa}$, is given equation

$$\hat{\kappa}_t(\theta) = \hat{\kappa}_Z \cdot \hat{Z}_t = -\frac{\hat{a}_t^{\Delta}(\theta)}{\overline{a}_a^{\Delta}(\theta)}. \tag{37}$$

Using $\overline{a}_a$ and $\hat{a}_t$, one can construct $\overline{a}_a^{\Delta}$ and $\hat{a}_t^{\Delta}$ and use this expression to compute $\hat{\kappa}_t$. While $\{\hat{\kappa}_t\}_t$ do not affect aggregate variables to the first order, they will play a role in characterizing second-order approximations.

3.4 Second-order approximations

We now extend our analysis of the HA to the second order. As with the first order, the statement of Lemma 2 is unaffected by heterogeneity as long as we include $\overline{\Omega}_{\sigma\sigma}$ in the last term on the right hand side of (14). Thus, similar to the first order, the directional derivatives that characterize second-order approximations remain unchanged. We proceed as in that section by differentiating G and constructing directional derivatives defined in that lemma. One obtains the following expressions

$$\mathsf{G}_Y \hat{Y}_{t,t+k} + \hat{\mathsf{G}}_{t,t+k} + \mathsf{G}_x \underbrace{\left(\left(\int \overline{x} d\Omega \right)_Z \cdot \hat{Z}_{t,t+k} + \left(\int \overline{x} d\Omega \right)_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k} \right) \right)}_{:= (\widehat{\int \overline{x} d\Omega})_{t,t+k}} = 0, \tag{38}$$

and

$$\mathsf{G}_Y \hat{Y}_{\sigma\sigma,t} + \mathsf{G}_x \underbrace{\left(\left(\int \overline{x} d\Omega \right)_{\sigma\sigma} + \left(\int \overline{x} d\Omega \right)_Z \cdot \hat{Z}_{\sigma\sigma,t} \right)}_{:= (\widehat{\int \overline{x} d\Omega})_{\sigma\sigma,t}} = 0, \tag{39}$$

where $\hat{Y}_{t,t+k}$ and $\hat{Y}_{\sigma\sigma,t}$ are as defined in Lemma 2.[16]

As with the first order, heterogeneity only adds one more term on the right hand sides of (38) and (39). Explicit calculations of these derivatives yields

$$\left(\widehat{\int \overline{x} d\Omega}\right)_{t,t+k} = \int \hat{x}_{t,t+k} d\Omega^* + \int \overline{x} d\hat{\Omega}_{t,t+k} + \int \hat{x}_{t+k} d\hat{\Omega}_t + \int \hat{x}_t d\hat{\Omega}_{t+k}, \tag{40}$$

$$\left(\widehat{\int \overline{x} d\Omega}\right)_{\sigma\sigma,t} = \int \hat{x}_{\sigma\sigma,t} d\Omega^* + \int \overline{x} d\hat{\Omega}_{\sigma\sigma,t}. \tag{41}$$

While equality (40) might seem obvious at first sight, showing it requires some care because of the kinks in policy functions. It can be written in this simple form because $\hat{x}_{t,t+k}$ is a generalized derivative that we characterized in equation (28).

We want to simplify integrals that appear in on the right hand side of equations (40) and (41). As in our first-order analysis, we first differentiate F twice to characterize the relationship between individual and aggregate variables to the second order.

Lemma 5. *(a). For any t, k*

$$\hat{x}_{t,t+k}(a,\theta) = \sum_{s=0}^{\infty} \mathsf{x}_s(a,\theta)\,\hat{Y}_{t+s,t+k+s} + \mathsf{x}_{t,t+k}(a,\theta), \tag{42}$$

where $\mathsf{x}_{t,t+k}(a,\theta) = \mathring{\mathsf{x}}_{t,t+k}(a,\theta) + \overline{x}_a^{\Delta}(\theta)\hat{\kappa}_t(\theta)\hat{\kappa}_{t+k}(\theta)\delta(a-\overline{\kappa}(\theta))$ with $\mathring{\mathsf{x}}_{t,t+k}$ solving a recursion

$$\mathring{\mathsf{x}}_{t,t+k}(a,\theta) = \left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\,\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{p}\right)^{-1}\left(\mathsf{F}_{t,t+k}(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\,\mathbb{E}_\varepsilon\left[\mathring{\mathsf{x}}_{t+1,t+k+1}|a,\theta\right]\right), \tag{43}$$

and $\mathsf{F}_{t,t+k}(a,\theta)$ combines known first-order interaction terms given explicitly in Appendix B.9.

(b). For any t

$$\hat{x}_{\sigma\sigma,t}(a,\theta) = \sum_{s=0}^{\infty} \mathsf{x}_s(a,\theta)\,\hat{Y}_{\sigma\sigma,t+s} + \mathsf{x}_{\sigma\sigma}(a,\theta), \tag{44}$$

where $\mathsf{x}_{\sigma\sigma}(a,\theta) = 0$ at the kinks $a = \overline{\kappa}(\theta)$ and for all other (a,θ) solves,

$$0 = \mathsf{F}_x(a,\theta)\,\mathsf{x}_{\sigma\sigma}(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\hat{x}_{0,0}|a,\theta\right] var(\mathcal{E}) + \mathbb{E}_\varepsilon\left[\mathsf{x}_{\sigma\sigma}|a,\theta\right] + \mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{p}\mathsf{x}_{\sigma\sigma}(a,\theta)\right). \tag{45}$$

Equation (44) shows that the relationship between $\hat{x}_{\sigma\sigma,t}$ and $\hat{Y}_{\sigma\sigma,t}$ is almost the same as between $\hat{x}_t$ and $\hat{Y}_t$ with an exception of an additional term $\mathsf{x}_{\sigma\sigma}$ which captures how agents react to risk holding aggregates fixed. Equation (42), which shows the relationship between $\hat{x}_{t,t+k}$ and $\hat{Y}_{t,t+k}$, is more involved. The second-order change in individual policy functions $\hat{x}_{t,t+k}$ consists of two terms: the first-order response to the second-order changes in the aggregates, captured by the infinite sum, and

[16] $\hat{\mathsf{G}}_{t,t+k}$ in (38) is defined analogously to $\hat{\mathsf{G}}_{t,t+k}$ in equation (15), adjusting for the fact that G is now a function of two arguments, Y and $\int x d\Omega$. In particular, let $\mathcal{Y} = \left[Y, \int x d\Omega\right]^{\mathrm{T}}$ and $\hat{\mathcal{Y}}_t = \left[\hat{Y}_t, \left(\int \overline{x} d\Omega\right)_Z \cdot \hat{Z}_t\right]^{\mathrm{T}}$. Then G is a (multi-dimensional) function of $\mathcal{Y}$, $G(\mathcal{Y})$, $\mathsf{G}_{\mathcal{Y}\mathcal{Y}}$ is its Hessian evaluated at the steady state $\mathcal{Y}$, and $\hat{\mathsf{G}}_{t,t+k} = \mathsf{G}_{\mathcal{Y}\mathcal{Y}} \cdot (\hat{\mathcal{Y}}_t, \hat{\mathcal{Y}}_{t+k})$.

second-order responses to the first-order interactions of shocks, captured by $\mathsf{x}_{t,t+k}$. Importantly, this lemma also provides an explicit formula for $\{\mathsf{x}_{t,t+k}\}_{t,k}$ exclusively in terms of objects from the first-order solution. The generalized function $\mathsf{x}_{t,t+k}$ consists of two parts: the classical derivative $\mathring{\mathsf{x}}$ that has form very similar to equation (33) in Lemma 3 and kink adjustments captured by the delta function.

The next step is to simplify the integrals that appear in (40) and (41) by differentiating LoM (27) twice. In Section 3.3 we showed that operators $\mathcal{M}$, $\mathcal{L}^{(a)}$, and $\mathcal{I}^{(a)}$ were central to describing the LoM of $\hat{\Omega}_t$. Modifications of the same three operators characterize the second-order approximation of the LoM. Let $\mathcal{L}^{(a)}_{Z,t}$ be the derivatives of $\mathcal{L}^{(a)}$ with respect to Z evaluated in direction $\hat{Z}_t$. Mathematically, it takes the same form as $\mathcal{L}^{(a)}$ except the function $\overline{a}_a$ in its definition is replaced with $\hat{a}_{aZ,t} := \overline{a}_{aZ} \cdot \hat{Z}_t$ Similarly, we use notations $\mathcal{L}^{(aa)}$, $\mathcal{L}^{(a,a)}$, etc to denote modifications of these operators where $\overline{a}_a$ is replaced with $\overline{a}_{aa}$ and $\overline{a}_a \overline{a}_a$ respectively. Analogous convention applies to $\mathcal{I}^{(a)}$. Finally, we use notation $y' \odot y''$ for two generalized functions y', y'' to denote their point-wise product.

Lemma 6. *(a). For all t, k*

$$\frac{d}{d\theta}\hat{\Omega}_{t+1,t+k+1} = \mathcal{L}^{(a)} \cdot \frac{d}{d\theta}\hat{\Omega}_{t,t+k} - \mathcal{M} \cdot \hat{a}_{t,t+k} + \frac{d}{da}\mathsf{c}_{t,t+k} - \mathsf{b}_{t,t+k}, \tag{46}$$

where $\mathsf{b}_{t,t+k}$ *and* $\mathsf{c}_{t,t+k}$ *satisfy*

$$\mathsf{b}_{t,t+k} = -\mathcal{L}^{(a)}_{Z,t} \cdot \frac{d}{d\theta}\hat{\Omega}_{t+k} - \mathcal{L}^{(a)}_{Z,t+k} \cdot \frac{d}{d\theta}\hat{\Omega}_t,$$

$$\mathsf{c}_{t,t+k} = \mathcal{M} \cdot (\hat{a}_t \odot \hat{a}_{t+k}) - \mathcal{L}^{(a)} \cdot \left(\frac{d}{d\theta}\hat{\Omega}_t \odot \hat{a}_{t+k}\right) - \mathcal{L}^{(a)} \cdot \left(\frac{d}{d\theta}\hat{\Omega}_{t+k} \odot \hat{a}_t\right).$$

(b). $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}$ *satisfies recursion (34) with* $\hat{a}_t = \mathsf{p}\hat{x}_t$ *being replaced with* $\hat{a}_{\sigma\sigma,t} = \mathsf{p}\hat{x}_{\sigma\sigma,t}$.

This lemma shows that the LoM for $\hat{\Omega}_{t+1,t+k+1}$ consists of two types of terms. The first type of term, captured by $\mathcal{M} \cdot \hat{a}_{t,t+k}$ and $\mathcal{L}^{(a)} \cdot \frac{d}{d\theta}\hat{\Omega}_{t,t+k}$, represent the first-order response of the LoM $\overline{\Omega}$ to the second-order changes in policy functions and the previous distribution. These have the same mathematical structure observed in Lemma 4^{HA}. The second types of term represent the second-order response of $\overline{\Omega}$ to the first-order changes, and are captured by $\mathsf{c}_{t,t+k}$, $\mathsf{b}_{t,t+k}$. The LoM for $\hat{\Omega}_{\sigma\sigma,t}$ is the same as for $\hat{\Omega}_t$ due to absence of first-order precautionary motives.

Lemmas 5 and 6 allow us to characterize the second-order derivative of the aggregation equations.

Corollary 2. *(a). For all t, k*

$$\left(\widehat{\int \overline{x} d\Omega}\right)_{t,t+k} = \sum_{s=0}^{\infty} \mathsf{J}_{t,s}\hat{Y}_{t+s,t+k+s} + \mathsf{H}_{t,t+k},$$

where $\{\mathsf{H}_{t,t+k}\}_{t,k}$ *is characterized by the following linear recursive system*

$$\mathsf{H}_{t,t+k} = \int \mathsf{x}_{t,t+k} d\Omega^* + \mathcal{I}^{(a)} \cdot \mathsf{B}_{t,t+k} + \mathcal{I}^{(aa)} \cdot \mathsf{C}_{t,t+k} - \mathcal{I}^{(a)}_{Z,t+k} \cdot \frac{d}{d\theta}\hat{\Omega}_t - \mathcal{I}^{(a)}_{Z,t} \cdot \frac{d}{d\theta}\hat{\Omega}_{t+k},$$

and recursions for $\mathsf{C}_{t,s}$ *and* $\mathsf{B}_{t,s}$ *with initial conditions* $\mathsf{C}_{0,k} = \mathsf{B}_{0,k} = 0$ *and*

$$\mathsf{C}_{t+1,t+k+1} = \mathcal{M} \cdot (\hat{a}_t \odot \hat{a}_{t+k}) - \mathcal{L}^{(a)} \cdot \left(\frac{d}{d\theta}\hat{\Omega}_t \odot \hat{a}_{t+k}\right) - \mathcal{L}^{(a)} \cdot \left(\frac{d}{d\theta}\hat{\Omega}_{t+k} \odot \hat{a}_t\right) + \mathcal{L}^{(a,a)} \cdot \mathsf{C}_{t,t+k},$$

$$\mathsf{B}_{t+1,t+k+1} = \mathcal{M} \cdot \mathsf{px}_{t,t+k} - \mathcal{L}^{(a)}_{Z,t} \cdot \frac{d}{d\theta}\hat{\Omega}_{t+k} - \mathcal{L}^{(a)}_{Z,t+k} \cdot \frac{d}{d\theta}\hat{\Omega}_t + \mathcal{L}^{(a)} \cdot \mathsf{B}_{t,t+k} + \mathcal{L}^{(aa)} \cdot \mathsf{C}_{t,t+k}.$$

(b). For all t,

$$\left(\widehat{\int \overline{x} d\Omega}\right)_{\sigma\sigma,t} = \sum_{s=0}^{\infty} \mathsf{J}_{t,s}\hat{Y}_{\sigma\sigma,s} + \mathsf{H}_{\sigma\sigma,t}$$

where $\{\mathsf{H}_{\sigma\sigma,t}\}_t$ *satisfies recursion* $\mathsf{H}_{\sigma\sigma,0} = \int \mathsf{x}_{\sigma\sigma} d\Omega^*$ *and* $\mathsf{H}_{\sigma\sigma,t} = \mathsf{H}_{\sigma\sigma,t-1} + \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M} \cdot \mathsf{px}_{\sigma\sigma}$.

Using this result, we obtain the analogue for Proposition 2 for HA economies

Proposition 2^{HA}. *For any* k, $\{\hat{X}_{t,t+k}\}_t$ *satisfies*

$$\mathsf{G}_Y \hat{Y}_{ZZ,t,t+k} + \hat{\mathsf{G}}_{t,t+k} + \mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s}\hat{Y}_{s,s+k} + \mathsf{G}_x \mathsf{H}_{t,t+k} = 0 \text{ for all } t, \tag{47}$$

with $\{\hat{Y}_{t,t+k}\}_t$ *defined in Proposition 2,* $\{\mathsf{H}_{t,t+k}\}_t$ *given in Corollary 2,* $\mathsf{P}\hat{X}_{-1,k-1} = 0$*, and* $\lim_{t\to\infty} \hat{X}_{t,t+k} = 0$.

$\{\hat{X}_{\sigma\sigma,t}\}_t$ *satisfies*

$$\mathsf{G}_Y \hat{Y}_{\sigma\sigma,t} + \mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s}\hat{Y}_{\sigma\sigma,s} + \mathsf{G}_x \mathsf{H}_{\sigma\sigma,t} = 0 \text{ for all } t, \tag{48}$$

with $\{\hat{Y}_{\sigma\sigma,t}\}_t$ *defined in Proposition 2,* $\{\mathsf{H}_{\sigma\sigma,t}\}_t$ *given in Corollary 2,* $\mathsf{P}\hat{X}_{\sigma\sigma,-1} = 0$ *and* $\lim_{t\to\infty} \hat{X}_{\sigma\sigma,t} - \hat{X}_{\sigma\sigma,t-1} = 0$.

It is instructive to compare Proposition 2^{HA} with both Proposition 2 and Proposition 1^{HA}. The only new terms that heterogeneity adds are $\{\mathsf{H}_{t,s}\}_{t,s}$ in equation (47) and $\{\mathsf{H}_{\sigma\sigma,t}\}_t$ in equation (48). Those terms have linear recursive mathematical structure similar to that of $\{\mathsf{J}_{t,s}\}_{t,s}$, which can be exploited to construct them quickly and efficiently. We discuss this in more details in the next section.

4 Numerical implementation

In this section, we show how to implement the formulas that we derived in the previous sections using the output of off-the-shelf methods to compute the steady state of HA economies without aggregate shocks. We use the endogenous gridpoint method of Carroll (2006) to find individual policy functions and store them using quadratic splines. Let N_X, N_Y and N_x be dimensions of vectors $\overline{X}$, $\overline{Y}$ and $\overline{x}(a,\theta)$. We use arrows to indicate numerical analogues of theoretical objects. Thus, $\overrightarrow{x}$ denotes the numerical analogue of $\overline{x}$.

We start with the description of the splines used to approximate the steady state policy functions. A standard implementation discretizes the space of (a,θ) as a "coarse" set of knot points on which individual optimization problem are solved and then uses a "fine" grid with many more grid points to approximate the distributions and aggregates. We use $j = 1, ..., N_{sp}$ to denote elements of the coarse grid and $i = 1, ..., N_\Omega$ to denote elements of the fine grid, with $(a, \theta)_{[j]}$ and $(a, \theta)_{[i]}$ denoting the value of the individual state corresponding to the j^{th} and i^{th} coarse and fine grid points.

Individual policy functions are stored as a vector of coefficients on a set of common basis functions. Let $\left\{\phi^j(\cdot, \cdot)\right\}_{j=1}^{N_{sp}}$ be a collection of basis functions or splines, where each ϕ^j is differentiable and maps from (a, θ) into $\mathbb{R}$. The function $x(a, \theta)$ is represented using a $N_x \times N_{sp}$ matrix $\overline{x}^\#$ of spline coefficients, and two matrices Φ and $\widetilde{\Phi}$ of dimensions $N_{sp} \times N_\Omega$ and $N_{sp} \times N_{sp}$, where $\Phi[j', i] = \phi^{j'}((a, \theta)_{[i]})$ and $\widetilde{\Phi}[j', j] = \phi^{j'}((a, \theta)_{[j]})$. The values of the policy functions on the fine grid are recovered as $\overrightarrow{x} = \overline{x}^\# \Phi$, so that the i^{th} column of matrix $\overline{x}^\# \Phi$, that we denote by $(\overline{x}^\# \Phi)[i]$, corresponds to $\overline{x}((a, \theta)_{[i]})$.

The algorithm also returns the sparse $N_\Omega \times N_\Omega$ transition probability matrix $\overrightarrow{\Lambda}$ on the fine grid and the invariant distribution $\overrightarrow{d\Omega}^*$ as the N_Ω dimensional vector that satisfies $\overrightarrow{d\Omega}^* = \overrightarrow{\Lambda}\overrightarrow{d\Omega}^*$. Idiosyncratic shocks are discretized as $\{\varepsilon_k\}_{k=1}^K$ that occur with probabilities $\{\mu_k\}_{k=1}^K$. Kinks in policy functions are stored as a subset of coarse grid points $\aleph$, where $j \in \aleph$ denotes the point $(\theta, a)_{[j]}$is just below the kink while $(\theta, a)_{[j+1]}$ is just above the kink.

We now describe how one can construct linear systems of equations described in Propositions 1^{HA} and 2^{HA} using these objects. We first start with an observation that splines make it very easy to compute various derivatives and expectations of individual policy functions that show up in several of our expressions. Let Φ_a be an $N_{sp} \times N_\Omega$ matrix with elements $\Phi_a[j', i] = \phi_a^{j'}((a, \theta)_{[i]})$, where $\phi_a^{j'}$ is the derivative of j'-th spline with respect to a. Then $\overrightarrow{x}_a = \overline{x}^\# \Phi_a$ recovers derivatives $\overline{x}_a$ on the fine grid. Similarly, by defining $N_{sp} \times N_{sp}$ matrices $\widetilde{\Phi}^e$ and $\widetilde{\Phi}_a^e$ with coefficients $\widetilde{\Phi}^e[j', j] = \sum_{k=1}^K \mu_k \phi^{j'}(\overline{a}(a, \theta)_{[j]}, \rho\theta_{[j]} + \varepsilon_k)$ and $\widetilde{\Phi}_a^e[j', j] = \sum_{k=1}^K \mu_k \phi_a^{j'}\left(\overline{a}\left((a, \theta)_{[j]}\right), \rho\theta_{[j]} + \varepsilon_k\right)$, we can recover $\mathbb{E}[\overline{x}|a, \theta]$ and $\mathbb{E}[\overline{x}_a|a, \theta]$ on the coarse grid as $\overline{x}^\# \widetilde{\Phi}^e$ and $\overline{x}^\# \widetilde{\Phi}_a^e$. The same observation applies to constructing expectations of second derivatives, such as $\mathbb{E}[\overline{x}_{aa}|a, \theta]$. An important observation here is that since kinked policy functions are approximated with smooth splines, no separate adjustments for the δ function that appears in the definition of $\overline{x}_{aa}$ is needed. For example, we can recover $\mathbb{E}[\overline{x}_{aa}|a, \theta]$ on a coarse grid as $\overline{x}^\# \widetilde{\Phi}_{aa}^e$, where $\widetilde{\Phi}_{aa}^e$ is defined as the second-order analogue of $\widetilde{\Phi}_a^e$.[17]

For our approximations, we pre-compute all basis matrices and store them as sparse matrices. Pre-computing basis matrices allows us to reduce all necessary calculations to matrix algebra without any further nonlinear function calls.

[17] Alternatively, these can be recovered by differentiating the F mapping with respect to a, finding spline coefficients $\overline{x}_a^\#$ and computing $\mathbb{E}[\overline{x}_{aa}|a, \theta]$ as $\overline{x}_a^\# \widetilde{\Phi}_a$. In the online supplementary material we discuss trade-off between different ways of computing these derivatives.

4.1 Numerical implementation of the first-order approximation

To solve the linear system (36), we need to construct G_x, G_Y and $\{\mathsf{J}_{t,s}\}_{t,s}$. The first two terms, G_x, G_Y are constructed by automatically differentiating the algebraic expression for G and evaluating that expression at the steady state. In order to construct $\{\mathsf{J}_{t,s}\}_{t,s}$, we need two sets of objects: operators $\mathcal{I}^{(a)}$, $\mathcal{L}^{(a)}$ and $\mathcal{M}$, and functions $\{\mathsf{x}_t\}_t$.

Numerical analogues of the three operators are constructed directly from their definitions as $\overrightarrow{\mathcal{I}}^{(a)}[:,i] = \overrightarrow{x}_a[i]$, $\overrightarrow{\mathcal{L}}^{(a)}[i',i] = \overrightarrow{\Lambda}[i',i]\overrightarrow{a}_a[i]$ and $\overrightarrow{\mathcal{M}}[i',i] = \overrightarrow{\Lambda}[i',i]\overrightarrow{d\Omega}^*[i]$. Matrices $\overrightarrow{\mathcal{L}}^{(a)}$ and $\overrightarrow{\mathcal{M}}$ are large (of dimension $N_\Omega \times N_\Omega$) but sparse. To construct $\{\mathsf{x}_t\}_t$ we use equations (32) and (33). We start by automatically differentiating F and calculating values of those derivatives in steady state on the same space of grid points used to compute individual policy functions. This produces arrays $\overrightarrow{\mathsf{F}}_x$, $\overrightarrow{\mathsf{F}}_{x^e}$, $\overrightarrow{\mathsf{F}}_Y$. For example, $\overrightarrow{\mathsf{F}}_x$ is an $N_x \times N_x \times N_{sp}$ array and $\overrightarrow{\mathsf{F}}_x[j]$ is a $N_x \times N_x$ matrix corresponding to $\mathsf{F}_x\left((a,\theta)_{[j]}\right)$. Using these arrays and equation (32), we construct the $N_x \times N_Y \times N_{sp}$ array $\overrightarrow{\mathsf{x}}_0$ with coefficients

$$\overrightarrow{\mathsf{x}}_0[j] = -\left(\overrightarrow{\mathsf{F}}_x[j] + \overrightarrow{\mathsf{F}}_{x^e}[j]\left(\mathsf{p}\overline{x}^{\#}\widetilde{\Phi}^e_a\right)[j]\right)^{-1}\overrightarrow{\mathsf{F}}_Y[j],$$

which is the numerical analogue of x_0 on the coarse grid. To convert it to the fine grid, one recovers the spline coefficients $\mathsf{x}_0^{\#} = \overrightarrow{\mathsf{x}}_0\widetilde{\Phi}^{-1}$ and then obtains the the fine grid analogue $\mathsf{x}_0^{\#}\Phi$. The rest of $\{\overrightarrow{\mathsf{x}}_s\}_{s>0}$ are constructed recursively from (33) using

$$\overrightarrow{x}_{s+1}[j] = -\left(\overrightarrow{\mathsf{F}}_x[j] + \mathsf{F}_{x^e}[j]\left(\mathsf{p}\overline{x}^{\#}\widetilde{\Phi}^e_a\right)[j]\right)^{-1}\overrightarrow{F}_{x^e}[j]\left(\mathsf{x}_s^{\#}\widetilde{\Phi}^e\right)[j]$$

with spline coefficients recovered as $\mathsf{x}_s^{\#} = \overrightarrow{\mathsf{x}}_s\widetilde{\Phi}^{-1}$. Once these $\{\overrightarrow{x}_s\}_s$ are constructed, we compute $\{\mathsf{J}_{t,s}\}_{t,s}$ using recursion

$$\mathsf{J}_{t,s} = \mathsf{J}_{t-1,s-1} + \overrightarrow{\mathcal{I}}^{(a)}\left(\overrightarrow{\mathcal{L}}^{(a)}\right)^{t-1}\overrightarrow{\mathcal{M}}\left(\mathsf{p}\mathsf{x}_s^{\#}\Phi\right)$$

with initial conditions $\mathsf{J}_{0,s} = \mathsf{x}_s^{\#}\Phi\overrightarrow{d\Omega}^*$ and $\mathsf{J}_{t,0} = 0$. This procedure recovers all objects necessary to invert (36) and obtain the first order solution $\{\hat{X}_t\}_{t=0}^T$.

Once $\{\hat{X}_t\}_{t=0}^T$ is computed, it is also straightforward to compute the responses of individual policies and the distribution of aggregate shocks. Using equation (31), $\hat{x}_t^{\#}$ can be computed using $\sum_{s=0}^{T-t}\mathsf{x}_s^{\#}\hat{Y}_{t+s}$ so that $(\hat{x}_t^{\#}\widetilde{\Phi})[j]$ and $(\hat{x}_t^{\#}\Phi)[i]$ correspond to $\hat{x}_t\left((a,\theta)_{[j]}\right)$ and $\hat{x}_t\left((a,\theta)_{[i]}\right)$ on the coarse and fine grids, respectively. Similarly, the numerical analogue of $\{\frac{d}{d\theta}\hat{\Omega}_t\}_t$ is a sequence of N_Ω dimensional vectors $\{\overrightarrow{\frac{d}{d\theta}\Omega}\}_t$ that is constructed recursively by $\overrightarrow{\frac{d}{d\theta}\Omega}_0 = \mathbf{0}$ and $\overrightarrow{\frac{d}{d\theta}\Omega}_t = \overrightarrow{\mathcal{L}}^{(a)}\overrightarrow{\frac{d}{d\theta}\Omega}_{t-1} - \overrightarrow{\mathcal{M}}\left(\mathsf{p}\hat{x}_t^{\#}\Phi\right)$. Finally, once $\hat{x}_t^{\#}$ is known, it is possible to construct $\hat{\kappa}_t$ for all $j \in \aleph$ as

$$\hat{\kappa}_t(\theta_{[j]}) = -\left(\mathsf{p}\overline{x}_a^{\Delta}(\theta_{[j]})\right)^{-1}\mathsf{p}\hat{x}_t^{\Delta}(\theta_{[j]}),$$

with $\overline{x}_a^{\Delta}(\theta_{[j]}) = (\overline{x}_a^{\#}\widetilde{\Phi})[j+1] - (\overline{x}_a^{\#}\widetilde{\Phi})[j]$ and $\hat{x}_t^{\Delta}(\theta_{[j]}) = (\hat{x}_t^{\#}\widetilde{\Phi})[j+1] - (\hat{x}_t^{\#}\widetilde{\Phi})[j]$.

4.2 Numerical implementation of the second-order approximation

Numerical implementation of the second order proceeds by direct analogy with the first order. For example, to obtain $\hat{\mathsf{G}}_{t,k}$ that appears in equation (47) we differentiate G twice and evaluate it at the steady state to get hessian $\mathsf{G}_{\mathcal{YY}}$ and then construct $\hat{\mathsf{G}}_{t,k}$ using its definition given in footnote 16. Operators $\mathcal{L}^{(a)}_{Z,t}$, $\mathcal{L}^{(aa)}$, etc are constructed just like their first-order analogues. For example, $\mathcal{L}^{(a)}_{Z,t}$ is represented by the $N_\Omega \times N_\Omega$ array $\overrightarrow{L}^{(a)}_{Z,t}$ with i element given by $\overrightarrow{\Lambda}[i', i](\mathsf{p}\hat{x}^{\#}_t \Phi_a)[i]$.

To construct $\{\overrightarrow{\mathsf{x}}_{t,t+k}\}_t$ we start with equation (43). Using backward induction and terminal condition $\overrightarrow{x}_{T+1,T+k+1} = 0$, one can construct the spline coefficients for the classical component $\{\mathring{\mathsf{x}}^{\#}_{t,t+k}\}^T_{t=1}$ in the same way we constructed $\{\mathsf{x}^{\#}_t\}^T_{t=1}$ for the first-order approximations. To adjust for the δ function part, rewrite the expression in Lemma 5 as

$$\mathsf{x}_{t.k}(z,\theta) = \mathring{\mathsf{x}}_{t+1,k+1} + \frac{d}{da}\left(\underbrace{\iota\left(a \geq \overline{\kappa}(\theta)\right)\overline{x}^{\Delta}(\theta)\hat{\kappa}_t(\theta)\hat{\kappa}_k(\theta)}_{\equiv \mathsf{x}^{\delta}_{t,k}(a,\theta)}\right).$$

The function $\mathsf{x}^{\delta}_{t,k}(z,\theta)$ is a step function that we can approximate with a spline with coefficients $\mathsf{x}^{\delta\#}_{t,k}$. We then recover $\overrightarrow{\mathsf{x}}_{t,k}$ as $\mathring{\mathsf{x}}^{\#}_{t,k}\Phi + \mathsf{x}^{\delta,\#}_{t,k}\Phi_a$.

These objects allow us to construct $\{\mathsf{H}_{t,t+k}\}_t$ using recursions in Corollary 2(a) and solve for $\{\hat{X}_{t,t+k}\}_t$ using equation (47). From this solution, we obtain $\hat{X}_{0,0}$ and $\hat{x}_{0,0}$ that are needed to find $\{\hat{X}_{\sigma\sigma,t}\}^T_{t=1}$. The terms $\mathsf{x}_{\sigma\sigma}$ still needs to be found in order to solve the system of equations in Proposition 2^{HA}. We use the linear system (45) to find $\mathsf{x}^{\#}_{\sigma\sigma}$ by evaluating (45) at each element of the coarse grid.

$$\overrightarrow{\mathsf{F}}_x[j]\mathsf{x}^{\#}_{\sigma\sigma}\widetilde{\Phi}[j] + \overrightarrow{\mathsf{F}}_{x^e}[j]\left(\hat{x}^{\#}_{0,0}\widetilde{\Phi}^e[j] + \mathsf{x}^{\delta,\#}_{0,0}\widetilde{\Phi}^e_a[j]\right)var\left(\mathcal{E}\right) + \overrightarrow{\mathsf{F}}_{x^e}[j]\overline{x}^{\#}_{\sigma\sigma}\widetilde{\Phi}^e[j] + \overrightarrow{\mathsf{F}}_{x^e}[j]\overline{x}^{\#}\widetilde{\Phi}^e_a[j]\left(\mathsf{p}\mathsf{x}^{\#}_{\sigma\sigma}\right)\widetilde{\Phi}[j] = 0.$$

This equation is linear in $\mathsf{x}^{\#}_{\sigma\sigma}$ and can be solved with a single linear operation. We then compute $\overrightarrow{x}_{\sigma\sigma} = \mathsf{x}^{\#}_{\sigma\sigma}\Phi$, construct $\{\mathsf{H}_{\sigma\sigma,t}\}^T_{t=1}$ using recursion in Corollary 2(b), and solve for $\{\hat{X}_{\sigma\sigma,t}\}^T_{t=1}$ using equation (47).

This procedure solves first- and second-order approximations of the stochastic economy. It only requires the user to supply an approximation of the steady state of deterministic economy and functional forms for G and F, the rest is computed automatically from those objects. With the exception of pre-computing basis matrices , all steps involve only linear algebra, which allows one to construct solutions to such economies quickly. We discuss the computational speed of our method in Section 7.

5 Extensions

We now discuss how our approach can be extended to three classes of problems that do not fit our description in Section 3: models with transition dynamics from some initial distribution to its long run steady state, models with stochastic volatility, and portfolio problems. The first class of problems emerges when one considers permanent shocks or policy changes that induce a transition to a new steady state, the second class of problems occurs frequently in studies of asset prices. Both of these extensions require minimal modifications of the procedure that we described in Section 3. The last extension is more substantial. Portfolio problems – models in which agents can invest in more than one asset with different risk characteristics – are commonplace. For example, the Krusell and Smith economy in which agents can borrow and lend from each other in addition to investing in risky capital falls into this category. Yet, solving such problems with standard perturbational techniques represents a substantial challenge. In such economies, the first-order approximation of equilibrium responses depend on the investment portfolios chosen by agents, yet the choice of the optimal portfolio depends on the second-order properties of the model such as risk premium. This breaks the convenient structure of recursive techniques under which one can find the n^{th} order of approximation from only knowing previous $n-1$ orders. Faced with this problem, much of quantitative HA macro literature simply ignores risk in characterizing agents' portfolio problems. We build on the ideas of Devereux and Sutherland (2011) and develop an approximation approach that allows one to handle portfolio problems in general HA settings.

5.1 Transition dynamics

The same techniques that we developed to characterize transition dynamics of the HA economy following an aggregate shock can be applied to study transition dynamics from any initial conditions to steady state. To illustrate that, consider economy as in Section 3 but suppose that the initial condition is given by $(0, A_{-1}, \Omega_0)$, where (A_{-1}, Ω_0) does not necessarily coincide with (A^*, Ω^*). Let $\hat{A}_0^{TD} = A_{-1} - A^*$, $\hat{\Omega}_0^{TD} = \Omega_0 - \Omega^*$ and $\hat{Z}_0^{TD} = \left[0, \hat{A}_0^{TD}, \hat{\Omega}_0^{TD}\right]^{\mathrm{T}}$. The following result extends Lemma 1 and Proposition 1^{HA} to handle transition dynamics:

Proposition 1^{TD}. *To the first-order approximation, X_t satisfies*

$$\mathbb{E}_0 X_t = \overline{X} + \hat{X}_t^{TD} + O\left(\left\|\mathcal{E}, \hat{Z}_0^{TD}\right\|^2\right).$$

Sequence $\{\hat{X}_t^{TD}\}_t$ satisfies

$$\mathsf{G}_Y \hat{Y}_t^{TD} + \mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \hat{Y}_s^{TD} + \mathsf{G}_x \mathsf{J}_t^{TD} = 0, \tag{49}$$

where $\hat{Y}_t^{TD} = \left[0, \mathsf{P}\hat{X}_{t-1}^{TD}, \hat{X}_t^{TD}, \hat{X}_{t+1}^{TD}\right]^{\mathrm{T}}$, $\mathsf{P}\hat{X}_{-1}^{TD} = \hat{A}_0$, and $\mathsf{J}_t^{TD} = \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \left(-\frac{d}{d\theta}\hat{\Omega}_0^{TD}\right)$.

Proposition 1^{TD} shows that to compute transition dynamics to the first order, one only needs to construct $\{\mathsf{J}_t^{TD}\}_t$ using the same operators $\mathcal{I}^{(a)}$ and $\mathcal{L}^{(a)}$ that are used to construct $\{\mathsf{J}_{t,s}\}_{t,s}$. These matrices capture the direct effect of the initial distribution $\hat{\Omega}_0$ on the transition dynamics of aggregate variables.

5.2 Stochastic volatility

Many applications that study financial markets or effects of government policies require the volatility of exogenous aggregate variables to be time-varying. A standard way to approximate such models is to consider third-order expansions (see, e.g., discussion in Fernández-Villaverde et al. (2011)). While it is possible to use a third-order extension of our techniques to model stochastic volatility, in this section we show a much simpler second-order approximation.

Suppose that stochastic process for Θ_t is given by (1) but $\mathcal{E}_t$ is not homoskedastic but rather follows the process

$$\mathcal{E}_t = \sqrt{1+\Upsilon_{t-1}}\mathcal{E}_{\Theta,t}, \tag{50}$$

$$\Upsilon_t = \rho_\Upsilon \Upsilon_{t-1} + \mathcal{E}_{\Upsilon,t}, \tag{51}$$

where $|\rho_\Upsilon| < 1$ and $\mathcal{E}_{\Theta,t}$ and $\mathcal{E}_{\Upsilon,t}$ are mean-zero i.i.d. variables with support of $\mathcal{E}_{\Upsilon,t}$ bounded so that Υ_t always remains greater than -1. The conditional volatility of aggregate innovations is stochastic and satisfies $var_{t-1}(\mathcal{E}_t) = (1+\Upsilon_{t-1})var(\mathcal{E}_{\Theta,t})$. This model collapses to that of Section 3 when Υ_t is a degenerate stochastic process, $\Upsilon_t \equiv 0$.

The state in the recursive representation now consists of a tuple $(\Upsilon, \Theta, A, \Omega)$. One way to approximate this economy is to scale both shocks $\mathcal{E}_{\Theta,t}$ and $\mathcal{E}_{\Upsilon,t}$ with σ and approximate the equilibrium around the deterministic point $(0, 0, A^*, \Omega^*)$. In order to capture time-varying volatility, this approach would indeed require using third-order approximations. Instead, a much faster and simpler method is to proceed as in Section 3 and scale only the combined shock $\mathcal{E}_t$ with σ, just as we did in equation (4). Since shocks $\mathcal{E}_{\Upsilon,t}$ and $\mathcal{E}_{\Theta,t}$ are not scaled with σ, Υ_t still satisfies (51) in the zeroth-order economy. Thus, our approximations are around $(\Upsilon, 0, A^*, \Omega^*)$ where Υ is stochastic.[18]

The convenience of this alternative perturbation can be seen from the next lemma that generalizes Lemma 2 to settings with stochastic volatility. Let $\boldsymbol{\mathcal{E}}^t = (\mathcal{E}^t, \mathcal{E}^t_\Upsilon)$.

Lemma 2^{SV}. *To the second-order approximation, X_t satisfies*

$$X_t\left(\boldsymbol{\mathcal{E}}^t\right) = \overline{X} + \sum_{s=0}^{t} \hat{X}_{t-s}\mathcal{E}_s + \frac{1}{2}\left(\sum_{s=0}^{t}\sum_{m=0}^{t} \hat{X}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m + \hat{X}_{\sigma\sigma,t}\right) + \frac{1}{2}\sum_{s=0}^{t} \hat{X}^{SV}_{\sigma\sigma,t-s}\mathcal{E}_{\Upsilon,s} + O\left(\|\mathcal{E}\|^3\right), \tag{52}$$

where sequences $\{\hat{X}_t\}_t$, $\{\hat{X}_{t,k}\}_{t,k}$, $\{\hat{X}_{\sigma\sigma,t}\}_t$ are the same ones as in Propositions 1^{HA} and 2^{HA}, and $\{\hat{X}^{SV}_{\sigma\sigma,t}\}_t$ is another sequence of directional derivatives.

[18] Our approach of scaling only the level of the innovation to TFP turns out to be similar to the approximation in Benigno et al. (2013) who study effects of time-varying variances in the context of a representative agent neoclassical growth model.

The expressions in the first line of equation (52) is identical to the characterization in models without volatility shocks. Thus, to approximate models with stochastic volatility one only needs to compute an additional sequence of directions $\{\hat{X}^{SV}_{\sigma\sigma,t}\}_t$ that capture how volatility shocks affect aggregate variables. Note that loadings on innovation to the volatility of aggregate shocks to the second order, $\sum_{s=0}^{t} \hat{X}^{SV}_{\sigma\sigma,t-s}\mathcal{E}_{\Upsilon,s}$, appear identically to the loading on innovations to the level of aggregate shocks to the first order, $\sum_{s=0}^{t} \hat{X}_{t-s}\mathcal{E}_s$, and has the same interpretation that we gave to the latter term following Lemma 1. The proof of the lemma shows that the simplicity of equation (52) is driven by two forces: our perturbation keeps $\mathcal{E}_{\Upsilon,t}$ non-degenerate in the steady state implies that second-order approximations already capture effects of volatility shocks; and functional form of (50) implies that those approximations are linear in $\{\mathcal{E}_{\Upsilon,t}\}_t$.

The next proposition combines and extends results from Lemma 5 and Proposition 2^{HA} to characterize $\{\hat{X}^{SV}_{\sigma\sigma,t}\}_t$.

Proposition 2^{SV}. *Let $\mathsf{x}^{SV}_{\sigma\sigma}$ be defined by*

$$\mathsf{F}_x(a,\theta)\,\mathsf{x}^{SV}_{\sigma\sigma}(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\hat{x}_{0,0}|a,\theta\right] var\left(\mathcal{E}\right) + \rho_\Upsilon \mathbb{E}\left[\mathsf{x}^{SV}_{\sigma\sigma}|a,\theta\right] + \mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{px}^{SV}_{\sigma\sigma}(a,\theta)\right) = 0, \tag{53}$$

with $\mathsf{x}^{SV}_{\sigma\sigma}(a,\theta) = 0$ at the kinks $a = \overline{\kappa}(\theta)$. Sequence $\{\hat{X}^{SV}_{\sigma\sigma,t}\}_t$ satisfies

$$\mathsf{G}_Y \hat{Y}^{SV}_{\sigma\sigma,t} + \mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \hat{Y}^{SV}_{\sigma\sigma,s} + \mathsf{G}_x \mathsf{H}^{SV}_{\sigma\sigma,t} = 0 \text{ for all } t, \tag{54}$$

where $\hat{Y}^{SV}_{\sigma\sigma,t} = \left[0, \mathsf{P}\hat{X}^{SV}_{\sigma\sigma,t-1}, \hat{X}^{SV}_{\sigma\sigma,t}, \hat{X}^{SV}_{\sigma\sigma,t+1} + \rho^t_\Upsilon \hat{X}_{0,0} var\left(\mathcal{E}\right)\right]^{\mathrm{T}}$, $\hat{X}^{SV}_{\sigma\sigma,-1} = 0$, and $\lim_{t\to\infty} \hat{X}^{SV}_{\sigma\sigma,t} = 0$. $\{\mathsf{H}^{SV}_{\sigma\sigma,t}\}_t$ is defined recursively as $\mathsf{H}^{SV}_{\sigma\sigma,0} = \int \mathsf{x}^{SV}_{\sigma\sigma} d\Omega^*$, $\mathsf{H}^{SV}_{\sigma\sigma,t} = \rho_\Upsilon \mathsf{H}^{SV}_{\sigma\sigma,t-1} + \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M} \cdot \mathsf{px}^{SV}_{\sigma\sigma}$.

Proposition 2^{SV} shows that one can find loadings $\{\hat{X}^{SV}_{\sigma\sigma,t}\}_t$ on the volatility shocks in the same way as we found precautionary motives $\{\hat{X}_{\sigma\sigma,t}\}_t$ in Section 3. First, one needs to find effects of volatility shocks on precautionary motives of individuals, captured by $\mathsf{x}^{SV}_{\sigma\sigma}$. This function is characterized by equation (53), which has almost identical structure to equation (45) that described $\mathsf{x}_{\sigma\sigma}$. Second, one constructs $\{\mathsf{H}^{SV}_{\sigma\sigma,t}\}_t$ by direct analogy with $\{\mathsf{H}_{\sigma\sigma,t}\}_t$ in Corollary 2. Finally, one solves (54) to find $\{\hat{X}^{SV}_{\sigma\sigma,t}\}_t$ in the same way we solve (48) to find $\{\hat{X}_{\sigma\sigma,t}\}_t$. Close parallels between solving for $\{\hat{X}^{SV}_{\sigma\sigma,t}\}_t$ and $\{\hat{X}_{\sigma\sigma,t}\}_t$ imply that it takes trivial amount of time to add stochastic volatility shocks.

5.3 Portfolio problems

To explain challenges that emerges from studying portfolio problems and our approach to overcome them, we start with the simplest version of such problem: the Krusell-Smith economy from Section 3.1 except suppose that households can also trade a one-period risk-free bond that is available in the zero net supply. Let R^f_{t-1} be the interest rate on this bond between periods $t-1$ and t, and $R^x_t = R_t - R^f_{t-1}$

be the excess return to capital. We use $a_{i,t}$ to denote the total wealth of agent i in period t and $k_{i,t}$ as the holdings in capital. Bond holdings are given by $a_{i,t} - k_{i,t}$. Assuming for concreteness that the borrowing constraint is on total asset holdings, the agents' optimization problem can be written as the choice of stochastic sequences $\{c_{i,t}, a_{i,t}, k_{i,t}\}_t$ to maximize their utility subject to the borrowing constraint $a_{i,t} \geq 0$ and the budget constraint

$$c_{i,t} + a_{i,t} - W_t \exp(\theta_{i,t}) - R^f_{t-1} a_{i,t-1} - R^x_t k_{i,t-1} = 0. \tag{55}$$

Agents' optimality conditions are represented by stochastic sequences $\{a_{i,t}, c_{i,t}, k_{i,t}, \zeta_{i,t}, \lambda_{i,t}\}_{i,t}$ that satisfy (55) and

$$R^f_{t-1} U_c(c_{i,t}) - \lambda_{i,t} = 0, \quad U_c(c_{i,t}) + \zeta_{i,t} - \beta \mathbb{E}_t \lambda_{i,t+1} = 0, \quad a_{i,t}\zeta_{i,t} = 0, \tag{56}$$

$$\mathbb{E}_{t-1}\left[\lambda_{i,t} R^x_t\right] = 0. \tag{57}$$

Market clearing conditions for aggregate variables $\left\{A_t, R^f_t, K_t, W_t, R_t\right\}_t$ remain (20) and (21) with the additional constraint that capital markets clear

$$K_t - \int k_{i,t-1} di = 0. \tag{58}$$

This is equivalent to imposing that the bond market clears: $\int (a_{i,t} - k_{i,t}) di = 0$, and equation (57) is the classical asset pricing equation that determines optimal portfolio allocations with $\lambda_{i,t}$ representing household i's stochastic discount factor.

One can immediately see from these equations that portfolio problems present a challenge to perturbational techniques. In the deterministic economy, $R^x_t = 0$ for all t. Therefore, while aggregate investments in capital and bonds as well as total assets $a_{i,t}$ are pinned down for all i, the allocation of those assets into capital and bonds, $k_{i,t}$ and $a_{i,t} - k_{i,t}$, is not determined for individual agents. At the same time, in the stochastic economy even the first-order approximation requires knowing the allocation of agents wealth into individual securities. As such, the individual state is now a triple (a, θ, k), and the distribution Ω is over (a, θ, k) with the aggregate state remaining $Z = [\Theta, A, \Omega]^T$.

We will keep $\overline{k}$ separate from $\overline{x}$ since $\overline{k}$ is undermined in the deterministic economy and thus the focus of our analysis. Otherwise the variables in $\overline{x}$ remain the same as in Section 3. Similarly, $\overline{X}$ contains one additional variable, R^f_t, which is also predetermined. Thus, the recursive representation consists of policy functions $\overline{x}(a, \theta, k, Z; \sigma)$, $\overline{k}(a, \theta, k, Z; \sigma)$, $\overline{X}(Z; \sigma)$. Vector $\overline{Y}(Z; \sigma)$ is as defined in Section 2. The deterministic steady remains the same as in Section 3, with the caveat that $\overline{k}(a, \theta, k)$ is undetermined in the deterministic economy, and thus, the joint distribution $\Omega^*(a, \theta, k)$ is also not pinned down. However, the marginal distribution $\Omega^*(a, \theta) = \int \Omega^*(a, \theta, k) dk$ is pinned down. Finding $\overline{k}(a, \theta, k)$ in the limit as $\sigma \to 0$ will be the key step.

Several elements of $\overline{x}$ and $\overline{X}$ will play an important role in our analysis. Let s be the selection matrix that picks out individual's stochastic discount factor out of vector $\overline{x}$, $\overline{\lambda} = \mathsf{s}\overline{x}$. Let R be the matrix that computes excess returns, $\overline{R}^x = \mathsf{R}\overline{Y}$,[19] and K be the selection matrix that selects the aggregate supply of risky assets, $\overline{K} = \mathsf{K}\overline{X}$. A general class of HA economies with portfolio problems can be represented by equation (26) for mapping G, equation (27) for $\overline{\Omega}$ (adjusted for the fact the distribution is defined over (a, θ, k) rather than (a, θ)) as well as the following three equations summarizing individual optimality conditions,

$$F\left(a, \theta, \overline{R}^x (Z; \sigma) k, \overline{x} (a, \theta, k, Z; \sigma), \mathbb{E}_{\varepsilon, \mathcal{E}} \overline{x}, \overline{Y} (Z; \sigma)\right) = 0 \text{ for all } (a, \theta, k, Z, \sigma), \tag{59}$$

$$\mathbb{E}_{\varepsilon, \mathcal{E}} \left[\overline{\lambda R}^x\right] = 0 \text{ for all } (a, \theta, k, Z, \sigma), \tag{60}$$

$$\int k d\Omega = \overline{K} (Z; \sigma) \text{ for all } Z, \sigma. \tag{61}$$

Equation (59) extends equation (22) to include the additional idiosyncratic state, with the restriction that the individual decisions depend on their portfolio choice solely through market value of their portfolio, $\overline{R}^x (Z; \sigma) k$. Equations (60) and (61) represent the first order conditions and market clearing of the portfolio choice $\overline{k}(a, \theta, k; \sigma)$.

Despite the addition of the portfolio choice, the statement of Lemma 1 remains unchanged. Thus, in order to find first-order approximation we need to solve for the vector of directional derivatives $\{\hat{X}_t\}_t$. Let $\hat{R}^x_0 = \mathsf{R}\hat{Y}_0$ be realized excess returns on assets. As in Section 2, $\overline{R}^x_{\sigma\sigma}$ is the second derivative of policy function of excess returns with respect to σ. We refer to $\overline{R}^x_{\sigma\sigma}$ as *risk premium.*

We proceed by first determining how individual choices $\hat{x}_t$ depend on the portfolio choice, whose analysis is very similar to that of Section 3. In particular, the portfolio problem analogue of Lemma 3 becomes

Lemma 3^{PP}. *$\hat{x}_0$ satisfies*

$$\hat{x}_0 (a, \theta, k) = \sum_{s=0}^{\infty} \mathsf{x}_s (a, \theta) \hat{Y}_s + \mathsf{r} (a, \theta) \hat{R}^x_0 k. \tag{62}$$

where

$$\mathsf{r} (a, \theta) = - \left(\mathsf{F}_x(a, \theta) + \mathsf{F}_{x^e}(a, \theta)\mathbb{E}_\varepsilon \left[\overline{x}_a | a, \theta\right] \mathsf{p}\right)^{-1} \mathsf{F}_k(a, \theta),$$

and F_k is the derivative of F with respect to $R^x k$ evaluated at the steady state. $\hat{x}_t (a, \theta, k)$ is independent of k and satisfies (31) for all $t > 0$. $\hat{R}^x_t := \mathsf{R}\hat{Y}_t = 0$ for all $t > 0$.

The intuition for Lemma 3^{PP} is straightforward. $\hat{R}^x_0 k$ is the realized return on the risky portfolio and r captures how this return affects individual policy functions at the time of the shock, $\hat{x}_0$. There

[19]For example, in the Krusell-Smith economy R would encode that $R^x_t = R_t - R^f_{t-1}$.

are no analogues of this term in the expressions for $\hat{x}_t$ for $t \neq 0$ since expected excess returns in the future are zero to the first order, $\hat{R}^x_t = 0$ for all $t > 0$.

The next step in finding aggregate responses is to characterize agents' portfolio choices, which were undetermined in the deterministic economy.

Lemma 7. *Agents' portfolios $\overline{k}$ satisfy*

$$\overline{k}(a,\theta,k) = \mathsf{v}_{\sigma\sigma}(a,\theta)\frac{\overline{R}^x_{\sigma\sigma}}{\left(\hat{R}^x_0\right)^2 var\left(\mathcal{E}\right)} + \sum_{s=0}^{\infty}\mathsf{v}_s(a,\theta)\frac{\hat{Y}_s}{\hat{R}^x_0}, \tag{63}$$

where

$$\mathsf{v}_{\sigma\sigma}(a,\theta) = -\frac{\mathbb{E}_\varepsilon\left[\mathsf{s}\overline{x}|a,\theta\right]}{\mathbb{E}_\varepsilon\left[\mathsf{sr}|a,\theta\right]}, \mathsf{v}_s(a,\theta) = -\frac{\mathbb{E}_\varepsilon\left[\mathsf{s}\mathsf{x}_s|a,\theta\right]}{\mathbb{E}_\varepsilon\left[\mathsf{sr}|a,\theta\right]},$$

Equation (63) derives expressions for the optimal portfolios for all agents in the limit as $\sigma \to 0$. These portfolios depend on asset's risk premium $\overline{R}^x_{\sigma\sigma}$ relative to a measure of volatility of its return $\left(\hat{R}^x_0\right)^2 var\left(\mathcal{E}\right)$, and on the relative exposures of aggregates and excess returns to shocks, $\hat{Y}_s/\hat{R}^x_0$.[20] These aggregate statistics, that characterize equilibrium properties of asset returns, are then weighted with individual weights $\mathsf{v}_{\sigma\sigma}$ and $\{\mathsf{v}_s\}_s$. These weights reflect individual attitudes towards risk and insurance that these assets offer, constraints that individuals face, etc. Importantly, individual weights can be computed directly from the steady state of deterministic economy much in the same way we computed $\{\mathsf{x}_s\}_s$ in Lemma 3. Thus, one can think of equation (63) as providing an explicit expression for $\overline{k}(a,\theta,k)$ in terms of yet unknown $\{\hat{Y}_t\}_t$ and risk premium $\overline{R}^x_{\sigma\sigma}$.

Lemmas 3^{PP} and 7 contain a couple of features that make analysis of the portfolio problem particularly tractable. Firstly, Lemma 3^{PP} implies that the portfolio choice only affects the agent's choices in the initial period, $\hat{x}_0$, and depends linearly on k. This implies that it is not necessary to know the full stationary distribution $\Omega^*(a,\theta,k)$, a single sufficient statistic will suffice: $k^*(a,\theta) := \int k\omega^*(a,\theta,k)dk$, which is the density weighted conditional mean of the capital holdings.[21] Secondly, $\overline{k}(a,\theta,k)$ is independent of k, which makes that sufficient statistic particularly easy to compute

$$k^*(a,\theta) = \int \overline{\Lambda}(a',\theta',a,\theta)\overline{k}(a,\theta)d\Omega^* = k^*_{\sigma\sigma}(a',\theta')\frac{\overline{R}^x_{\sigma\sigma}}{\left(\hat{R}^x_0\right)^2 var\left(\mathcal{E}\right)} + \sum_s k^*_s(a',\theta')\frac{\hat{Y}_s}{\hat{R}^x_0}$$

[20] Equation (63) is a special, two-asset case of a general formula for portfolio choice over multiple assets that can be written as

$$\overline{k}(a,\theta_-)^{\intercal} = \left[\mathsf{v}_{\sigma\sigma}(a,\theta_-)\overline{R}^x_{\sigma\sigma} + \sum_{s=0}^{\infty}\mathsf{v}_s(a,\theta_-)cov\left(\hat{Y}_s,\hat{R}^x_0\right)^{-1}\right]cov\left(\hat{R}^x_0,\hat{R}^x_0\right)^{-1}$$

and that is familiar from the portfolio theory (see, e.g., Viceira (2001)). See Appendix D.3 for details.

[21] To understand why this is relevant, consider the term $\int \hat{x}_0 d\Omega^*$. Lemma 3^{PP} implies that $\hat{x}_0$ is linear in k and, hence,

$$\int \hat{x}_0 d\Omega^* = \sum_s \int \mathsf{x}_s d\Omega^* \hat{Y}_s + \int \mathsf{r}k^* dad\theta \hat{R}^x_0.$$

where $k^*_{\sigma\sigma} := \mathcal{M} \cdot \mathsf{v}_{\sigma\sigma}$ and $k^*_s := \mathcal{M} \cdot \mathsf{v}_s$. Combining the definition $k^*(a,\theta)$ with equation (61), that gives asset supply, we obtain the equilibrium relationship between an asset's risk premium and the first-order behavior of aggregate variables:

$$\frac{\overline{R}^x_{\sigma\sigma}}{\left(\hat{R}^x_0\right)^2 var\left(\mathcal{E}\right)} = \frac{\overline{K}}{\mathsf{V}_{\sigma\sigma}} - \sum_{s=0}^{\infty} \frac{\mathsf{V}_s}{\mathsf{V}_{\sigma\sigma}} \frac{\hat{Y}_s}{\hat{R}^x_0}, \tag{64}$$

where $\mathsf{V}_{\sigma\sigma} = \int k^*_{\sigma\sigma}(a,\theta) da d\theta$, $\mathsf{V}_s = \int k^*_{\sigma\sigma}(a,\theta) da d\theta$.

We can use these insights to adjust Proposition 2^{HA} to account for the $\mathsf{r}\left(a,\theta\right)\hat{R}^x_0 k$ in Lemma 3^{PP}. To state our main result succinctly, define operators $\{\mathcal{N}_t\}_t$ that return

$$\begin{aligned}
\mathcal{N}_0 \cdot y &= \int \mathsf{r}(a,\theta) y(a,\theta) da d\theta, \\
\mathcal{N}_t \cdot y &= \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M}^{PP} \cdot y,
\end{aligned}$$

where $\mathcal{I}^{(a)}$ and $\mathcal{L}^{(a)}$ are the same operators we constructed for Lemma 4^{HA} and $\mathcal{M}^{PP}$ is defined by

$$\left(\mathcal{M}^{PP} \cdot y\right)\langle a', \theta' \rangle := \int \overline{\Lambda}(a', \theta', a, \theta) \mathsf{pr}(a,\theta) y\left(a,\theta\right) da d\theta,$$

This allows us to provide the following characterization of the first-order approximation

Proposition 1^{PP}. *$\{\hat{X}_t\}_t$ are the solution to (64) and*

$$\mathsf{G}_Y \hat{Y}_t + \mathsf{G}_x \sum_{s=0}^{\infty} \left(\mathsf{J}_{t,s} + \mathsf{J}^{PP}_{t,s}\right) \hat{Y}_s = 0 \text{ for all } t \tag{65}$$

where $\{\hat{Y}_t\}_t$ as defined in Proposition 1, $\mathsf{R}\hat{Y}_t = 0$ for $t \geq 1$, and $\mathsf{J}^{PP}_{t,s} := \mathcal{N}_t \cdot k^*_s + \frac{\mathcal{N}_t \cdot k^*_{\sigma\sigma}}{\mathsf{V}_{\sigma\sigma}}\left(-\mathsf{V}_s + 1_{s=0}\overline{K}\mathsf{R}\right)$.

Relative to the economy posed in Section 3, the portfolio problem modifies the matrix $\mathsf{J}_{t,s}$ by $\mathsf{J}^{PP}_{t,s}$ with $\mathsf{J}^{PP}_{t,s}$ being constructed via simple linear operators. The terms in $\mathsf{J}^{PP}_{t,s}$ capture how changing the response of aggregates, $\hat{X}_s$, alters the portfolio decision and feeds back into individual choices. Once $\mathsf{J}^{PP}_{t,s}$ is known, we can solve for $\hat{X}_s$ in the same manner as Section 3, and construct the risk premium from (64).[22]

6 Comparison to literature

Our approach builds on the perturbational techniques in the spirit of Judd (1998) and Schmitt-Grohé and Uribe (2004) originally developed to study dynamic representative agent models. The key difficulty

[22] With multiple shocks (65) becomes a non-linear equation, but remains linear in $\hat{X}_t$ conditional on $\hat{R}^x_0$. This presents a simple procedure for finding the first order equilibrium: guess the value of $\hat{R}^x_0$ and hold it fixed, solve this linear system of equations for $\{\hat{X}_t\}_t$, check whether this solution has $\mathsf{R}\hat{Y}_0$ that is consistent with our guess for $\hat{R}^x_0$, and update this guess if necessary.

It is also possible to impose an additional short-selling constraint $k_{i,t} \geq 0$ by guessing a function $\bar{\iota}^{PP}(a,\theta)$ which is 0 if the agents are on that constraint and 1 otherwise. Conditional on $\bar{\iota}^{PP}$, k^* can be constructed from $k^*_s = \mathcal{M} \cdot (\bar{\iota}^{PP} \odot \mathsf{v}_s)$ and $k^*_{\sigma\sigma} = \mathcal{M} \cdot \left(\bar{\iota}^{PP} \odot \mathsf{v}_{\sigma\sigma}\right)$ and then $\{\hat{X}_t\}_t$ can be solved from Proposition 1^{PP}. The implied portfolio choices $\overline{k}(a,\theta)$ can be used verify if the guess of $\bar{\iota}^{PP}$ is correct and update this guess as necessary.

in extending them to HA environments lies in the fact that derivatives of policy functions with respect to the aggregate state (captured by $\overline{X}_Z$, $\overline{X}_{ZZ}$, etc in our notation) are intractably large objects. The seminal paper by Reiter (2009) takes a step to overcome this hurdle by discretizing the state space and the transition probability kernel (using the so-called "histogram method", see also Young (2010)). To obtain first-order approximations, this method requires solving large quadratic matrix equations and it has proved to be too slow and imprecise in many standard HA environments.[23]

One strand of literature, originally proposed by Boppart et al. (2018) and then significantly developed by an important paper by Auclert et al. (2021), abandons the state-space representation used in Reiter (2009) and subsequent literature building on his ideas, and works with the sequence-space formulation of the problem. The key observation for that approach is that the first-order impulse responses of the stochastic economy can be fully constructed from deterministic responses to MIT shocks, and that these responses can be recovered numerically fairly easily from the sequence problem. Auclert et al. (2021) show that this can be done very fast as those impulse responses solve a linear system of equations which coefficients can be constructed using linear recursive equations.

Our approach combines insights from both strands of the literature and also introduces new ideas, such as using directional derivatives and generalized functions to characterize equilibrium approximations analytically. This allows us to improve on the computational speed of Auclert et al. (2021) method, and to have our approach scalable to second- and higher-orders of approximation, which is one of the key features of classical perturbational techniques a-la Judd (1998) and Schmitt-Grohé and Uribe (2004) but not of Reiter (2009) and papers building on it.

One distinction of our approach is that we start with the theoretical distribution and its LoM and derive exact analytical expressions for approximations of various orders; numerical values of those expressions are then computed using appropriate discretization. This contrast with papers following the Reiter (2009) tradition that start with an approximate (i.e., already discretized) distribution and a transition probability matrix before further approximating with respect to aggregate shocks. There is no guarantee that this latter approach would correctly recover aggregate responses beyond the first order. In particular, we show in Appendix C that the second-order approximation of the transition probability obtained under histogram method generically misses some of the second-order order terms and does not converge to the correct second-order expressions even as the grid size shrinks to zero.

[23]The key issues are both the time taken to compute and space needed to store those derivatives. This is most clear from Reiter's implementation of the Krusell and Smith model that is solved by discretizing the θ process and using a histogram to store the distribution Ω. If we were to follow the standard convention of using between 1000-5000 points per θ for the histogram, and use 10 points for the shocks, the size of the histogram $N_\Omega \sim 10^4$. This means that $\overline{\Omega}_Z \sim 10^8$ and $\overline{\Omega}_{ZZ} \sim 10^{16}$ entries. Assuming that 4 bytes (float) are required to store an entry, this would mean that one needs 450 megabytes of RAM to store the first derivative and 4 terabyte of RAM to store the second-order derivative, which is clearly outside the scope of the current computing architectures. The general argument also applies to variants of Reiter (2009) such as Bayer et al. (2022); Ahn et al. (2018); Childers (2018); Winberry (2018); Gornemann et al. (2021); Reiter (2023) who use a variety of model reduction techniques to reduce the dimension of histogram and the resulting tradeoff between speed and precision depends on the details of the method and the application studied.

The intuition for this results is that the histogram method locally linearizes the LoM for the aggregate distribution, which misses terms capturing second-order responses of the LoM to the first-order changes in policy functions.[24]

Our description of approximations as a sequence of values of derivatives such as $\{\hat{X}_t\}_t$, $\{\hat{X}_{t,s}\}_{t,s}$, etc is related to the literature the uses sequence-space formulation of the problem. One can show that to the first order, our approach is equivalent to that of Auclert et al. (2021) in the sense that as the grid size of their approximations goes to zero, the linear system of equations they use to describe approximations converges to our system (36). Despite this equivalence, using state-space representation has advantages even to the first order, as it allows us to derive analytically and then construct recursively coefficients $\mathsf{x}_s = \partial x_0 / \partial X_t$ in Lemma 3. In contrast, the sequence-space approach finds $\{\partial x_0 / \partial X_t\}_t$ using numerical differentiation of the (truncated) infinite system of equations (22). This process is both slower and less stable numerically. State space representation also significantly simplifies and speeds up computation of first-order transition dynamics, as in Section 5.1.[25]

The bigger advantage of using state-space representation over the sequence space approaches in the spirit of Boppart et al. (2018) and Auclert et al. (2021) is that it does not restrict us to only the first-order approximations and applies higher orders as well as to the models with portfolio choices. For instance, by explicitly specifying directions that characterize the effects of persistent risk $\{\hat{Z}_{\sigma\sigma,t}\}_t$, our approach can incorporate risk and go beyond MIT shocks. This is imperative for questions such as finding first-order impulse responses in models with portfolio choice, understanding effect of risk or welfare costs of aggregate shocks, and studying trade-offs involved in designing macroeconomic stabilization policies.

Our paper is also related to the approximation methods in Mertens and Judd (2018) and Bhandari et al. (2021). Like us, those authors use perturbational methods to derive analytically various orders of approximations of equilibrium in HA economy, and then find those expressions numerically. Their approximation scales both aggregate and idiosyncratic shocks and it is not applicable to models in which policy functions have kinks, for example due to the occasionally binding borrowing constraints. Mertens and Judd (2018) additionally approximate around an allocation in which all agents are identical. Our approach instead approximates only with respect to aggregate shocks. This improves the approximation precision, since our approach remains global with respect to idiosyncratic shocks, and allows study of economies in which policy functions have kinks. It also makes analytical characterization of approximation terms significantly more challenging. Deriving those analytical expressions to build

[24]Most papers following Reiter (2009) (with a few exceptions such as Gornemann et al. (2021) and Reiter (2023)) utilize only first-order approximations in their analysis, and therefore their conclusions remain unaffected by this observation.

[25]It should be noted that very careful attention has to be paid to those numeric derivatives in order to ensure that they are accurate, (See appendix C.1 of Auclert et al. for details) and the speed is often limited by efficiency of the global transition code. These numerical issues would be amplified with a second-order approximation as calculating second derivatives are more prone to numerical error. By giving explicit expressions for these second derivatives in terms of derivatives of F and G we sidestep these issues.

the approximations when policy functions are not differentiable is one of the key contributions of this paper.

Our approach, like all perturbational methods, is local as it seeks to find equilibrium dynamics when aggregate shocks are small and the economy is near its steady state. Our goal is to preserve key advantages of these methods – computational speed, simplicity, and flexibility – in HA settings. There exists a complementary strand of literature that aims to develop global methods. Such methods can be used to find equilibria without requiring them to be nearby any specific economy but they tend to be slower, harder to use, and often need to be tailored to the specific economic environment.[26]

The class of economies that we consider in this paper is discrete-time infinite horizon HA models with distributional states. There is a parallel literature that studies continuous-time versions of these economies. See, for instance, Kaplan et al. (2018), Achdou et al. (2020), Ahn et al. (2018) in the context of consumption-savings models; Alvarez and Lippi (2022) in the context of price-setting models; and Bigio et al. (2023) for an application to public debt maturity. In related work, Bilal (2023) and Alvarez et al. (2023) use mean field game techniques to construct approximations with aggregate shocks in these class of models. Their work shares with us the use of linear operators over infinite-dimensional spaces to analytically characterize the exact derivatives. We view their approaches as complementary since the mathematics underlying the approximations is quite different and the relative advantages of discrete vs continuous time vary by application. Those papers do not consider economies in which policy functions have kinks that are functions of endogenous states, or settings with heteroskedastic shocks or portfolio problems.

7 Numerical results

In this section, we apply our algorithm to calibrated versions of the Krusell and Smith (1998) model. First, we use the calibrated model to report diagnostics such as speed and accuracy and compare them to alternative methods. Second, we use extensions of the Krusell and Smith model to study several applications that illustrate the usefulness of our methods over and above what can be achieved with existing approaches.

7.1 Baseline model

Our baseline model extends the Krusell and Smith framework of Section 3 to include capital adjustment costs. This allows the model to generate adequately volatile returns to holding risky capital that is useful for some of our applications. To enable convenient aggregation, we introduce a competitive mutual funds

[26]Krusell and Smith (1998) solve their economy using a global method that proved difficult to extend to general HA settings. Some recent work extends global solution methods to more complex environments using machine learning techniques. See Maliar et al. (2021), Kahou et al. (2021), Childers et al. (2022), and Han et al. (2021) for details.

sector whose shares are owned and traded by households in the baseline.[27] The household's budget constraint is modified to

$$c_{i,t} + k_{i,t} = W_t \exp\{\theta_{i,t}\} + R_t k_{i,t-1},$$

where $k_{i,t} \geq 0$ now is the date t wealth of the household. The mutual fund gathers rental income from the corporate sector, owns and invests in physical capital subject to a convex adjustment of the form

$$\phi(I_t, K_t) = \frac{\phi}{2}\left(\frac{I_t}{K_t} - \delta\right)^2 K_t, \quad K_{t+1} = (1-\delta) K_t + I_t.$$

In the online supplementary materials Section B.3, we show that the competitive equilibrium is given by the equations (19) and (20) as before and a modified version of (21),

$$R_t = \frac{(1-\alpha)\exp(\Theta_t) K_t^{\alpha} - I_t - \frac{\phi}{2}\left(\frac{I_t}{K_t} - \delta\right)^2 K_t + Q_t K_{t+1}}{Q_{t-1} K_t}, \quad Q_t = 1 + \phi\left(\frac{I_t}{K_t} - \delta\right), \quad \int k_{i,t} di = Q_t K_{t+1}. \tag{66}$$

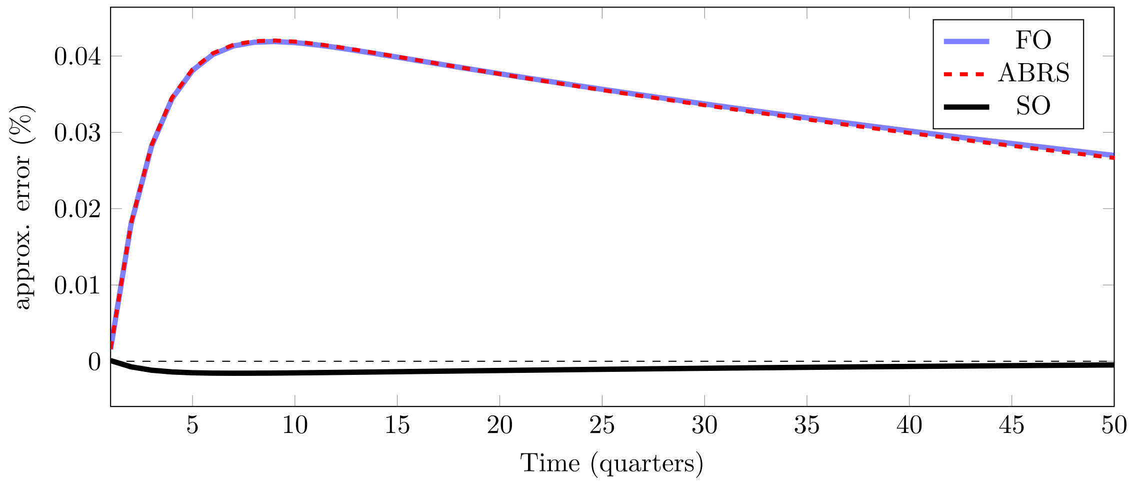
Calibration To calibrate our model, we set the period length to one quarter. The parameter α is set to 0.36 to target the capital share of income. We use an isoelastic period utility $U(c) = \frac{c^{1-\gamma}}{1-\gamma}$ and set the risk aversion parameter γ to equal 5. The adjustment cost parameter ϕ is calibrated to match a 3% standard deviation of un-leveraged quarterly returns to equity. For the parameters governing the aggregate and idiosyncratic labor productivity in (1) and (24), we choose values used by Auclert et al. (2021). The calibrated parameters are summarized in Table 2 in Appendix E.1. We solve the non-stochastic steady state policy functions using an endogenous grid method after discretizing the productivity with $N_\epsilon = 7$ and asset grid $N_z = 120$. We use $N_\Omega = 1000 \times 7$ points to store the distribution.

Accuracy We test the accuracy of ours and alternative methods by studying the response to a one-time, one standard deviation positive shock to TFP which can be solved non-linearly and compared to the approximations $\hat{X}_t$ under alternative perturbational methods. In the right panel of Figure 1, we plot the % error in the capital stock. For comparison purposes we show errors using our approach described in Section 3.3 and the Sequence Space Jacobian approach of Auclert et al. (2021). As would be anticipated by Figure 1, both approaches have roughly the same error to first order, with the maximal error being on the order of 0.04% of the capital stock. At higher orders, our approach has errors which remain very small over time.

Speed We now use the baseline calibration to simulate policy functions and compare the time taken under our method to alternatives. To compute the first- and second-order terms, we implement the

[27] We later study an extension in which households directly hold capital and bonds.

Figure 1: APPROXIMATION ERRORS



Notes: The figure plots $\frac{\hat{K}_t - \tilde{K}_t}{\tilde{K}_t}$ for $\mathcal{E}^t = (1, 0, 0, \ldots, 0)$ where the path for $\hat{K}_t$ is obtained using perturbational approaches and the path for $\tilde{K}_t$ is obtained using nonlinear methods. The solid blue line "FO" is the first-order approximation from Section 3.3. The dotted red line "ABRS" is the first-order approximation using the sequence-space method of Auclert et al. (2021). The solid black line "SO" is the second-order approximation from Section 3.4.

steps detailed in Section 4.1 and 4.2. In Table 1, we report total time taken to compute those terms and break up the time by each step stage of the algorithm. The timings for the first-order approximation are reported in the first two columns of the table and the timings for steps to compute the second-order are reported rest of the columns.

Table 1: COMPUTATIONAL SPEED: FIRST AND SECOND ORDER

First Order		Second Order		
Step	Time	Step	Time (ZZ)	Time$(\sigma\sigma)$
		Additional First-Order Terms	0.63s	
Lemma 3 Terms	0.11s	Lemma 5 Terms	1.28s	0.25s
Lemma 4^{HA} Terms	0.02s	Lemma 6 Terms	0.27s	
Corollary 1 Terms	0.21s	Corollary 2 Terms	0.18s	0.00s
Proposition 1^{HA} Terms	0.10s	Proposition 2^{HA} Terms	0.20s	0.03s
Total	0.44s		2.57s	0.28s
ABRS	0.51s			

Notes: The table summarizes the time taken to compute the first and second-order terms using Section 3.3 and 3.4. The last row in the bottom "ABRS" refers to the time taken to approximate our calibrated model using using the sequence-space method of Auclert et al. (2021) keeping the same size for the grids and length of trucation horizon. All numbers are reported using a 20 core M1 ultra mac studio.

All told, once the steady state has been computed, our algorithm takes 0.44 seconds to solve for the $\left\{\hat{X}_t\right\}_t$ terms with roughly half the time spent computing the $\{\mathsf{J}_{t,s}\}_{t,s}$ terms. As Lemma 1 highlights,

$\left\{\hat{X}_t\right\}_t$ are all that is needed to simulate the path of aggregates and to compute ergodic moments from the first-order approximation. The other first-order terms $\{\hat{x}_t\}_t$ and $\left\{\hat{\Omega}_t\right\}_t$, are required for the second-order approximation and take an additional 0.6 seconds to compute. We compare this to our own implementation of the Sequence Space Jacobian of Auclert et al. (2021) which takes approximately 0.51 seconds to compute the equivalent on the $\left\{\hat{X}_t\right\}_t$.[28] As mentioned in Section 6, the small difference arises because we use expressions (32)–(33) to compute the exact derivatives while Auclert et al. (2021) relies on numerical differentiation.

The addition time to compute the second-order approximation is broken out in the last two columns of Table 1. As highlighted in Section 3.4 there are two additional types of terms in the second-order approximation: the curvature terms $\left\{\hat{X}_{t,k}\right\}_{t,k}$, and precautionary terms $\left\{\hat{X}_{\sigma\sigma,t}\right\}_t$. As they follow the same mathematical structure, we break out the computational time separately for both types. The curvature terms take 2.57 seconds to compute[29] while the risk adjustment terms take about 0.3 seconds. The vast majority of the computational time for the curvature terms is spent on Lemma 5 and Proposition 2^{HA} which is a result of a large number of quadratic forms required to compute the $\{\mathsf{x}_{t,k}(z,\theta)\}_{t,k}$ and $\{\mathsf{G}_{\Theta,t,k}\}_{t,k}$ terms. All combined, computing the second-order approximation requires an additional 3 seconds relative to the first-order approximation.

7.2 Applications

In this section we study four applications that highlight the usefulness our method for heterogeneous agent models. As a reference, we also compare results to a version of the model without heterogeneity (RA) that is calibrated to hit the same aggregate moments as the baseline with heterogeneity.

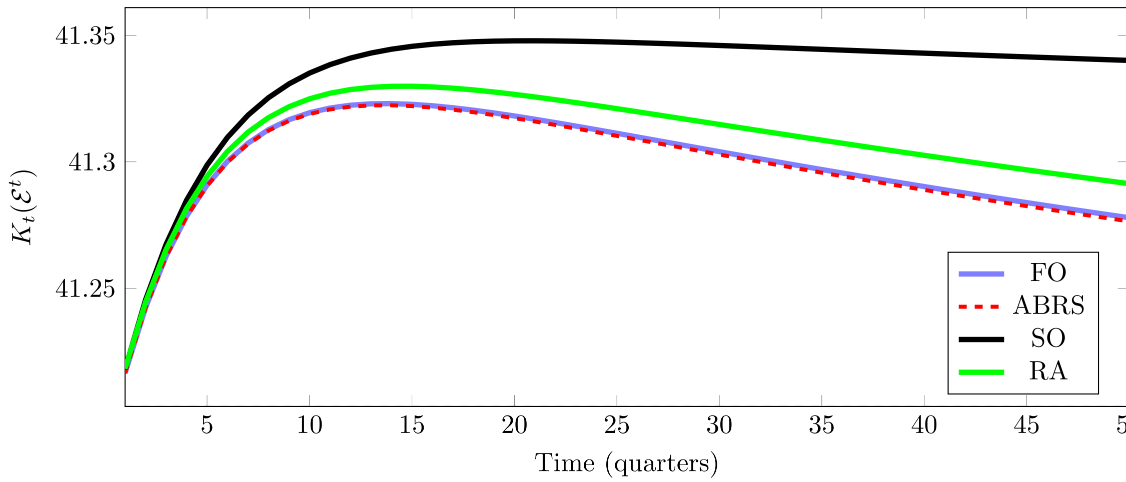
7.2.1 Simulations

Our first application uses simulations to asses the magnitude and sources of nonlinearities in the Krusell and Smith model under the baseline calibration. To do that, we use Lemma 2 to construct the second-order approximation for a given path of $\mathcal{E}^t$ and inspect various terms. In Figure 2, we plot simulations for one-time impulse at date $t = 0$, that is, $\mathcal{E}^t = (1, 0, 0, \ldots, 0)$. The plots show that second-order path with heterogeneity is quite different from the first-order path as well as the RA counterpart emphasizing

[28] Auclert et al. report that it takes about 0.1s to construct IRFs in the Krusell and Smith economy. Our implementation of their algorithm reported in Table 1 takes longer because we use a different baseline calibration that has capital adjustment costs, more points on the size of the asset grid, and a longer truncation horizon that is better suited for our applications. When we reset the parameters and tolerance to Auclert et al. choices, our method takes 0.1s and our implementation of their algorithm takes 0.121s to compute the first-order code.

[29] Here we report only the time required to compute that $\bar{X}_{ZZ,t,t}$ terms. We do this for two reasons. Firstly, for most ergodic moments only the $\bar{X}_{ZZ,t,t}$ are required. Secondly, computing the addition $\bar{X}_{ZZ,t,t+i}$ terms are trivially parallelizable for each i so, with enough processors, computing all the $\bar{X}_{ZZ,t,k}$ terms would not require any additional time.

Figure 2: SIMULATED PATHS: $K_t\left(\mathcal{E}^t\right)$.



Notes: The figure plots the simulated path for aggregate capital $K_t\left(\mathcal{E}^t\right)$ for a sequence of shocks $\mathcal{E}^t = (1, 0, 0, \ldots)$. The solid blue line "FO" is the first-order approximation from Section 3.3. The dotted red line "ABRS" is the first-order approximation using the sequence-space method of Auclert et al. (2021). The solid black line "SO" is the second-order approximation from Section 3.4. The solid green "RA" line plots the simulation using representative agent counterpart targeted to the same set of aggregate moments.

the importance of heterogeneity and nonlinear effects.[30]

From Lemma 2, these differences between the first- and second-order can come either due to $\hat{K}_{0,0}$ terms or $\left\{\hat{K}_{\sigma\sigma,t}\right\}_t$ terms. The $\hat{K}_{0,0}$ term captures nonlinearities in capital demand due to curvature embedded in choice of technologies–production and investment. In our calibration this term turns out to be small which is consistent with the intuition that aggregate policy functions are approximately linear in the neoclassical growth models. The $\left\{\hat{K}_{\sigma\sigma,t}\right\}_t$ terms aggregates precautionary behavior in capital supply by households. The strength of the precautionary motives is determined by household's risk aversion, volatility of aggregate shocks, and mass of households near the borrowing constraint. This precautionary motive accounts for virtually all of the difference between the FO and SO lines in Figure 2.

So far we studied responses to a one-time transitory shock after which aggregate variables converge back to their values in the non-stochatic steady state. Often of-interest are experiments involving permanent changes in technology or regulatory parameters. With a permanent change, the endogenous state Z_t converges to a new steady-state. The extension of our method in Section 5.1 can handle such thought experiments. To illustrate this, consider a one-time-forever increase in aggregate TFP. In Appendix E.2, we show how to apply the formulas from Lemma 1^{TD} to simulate the path for aggregate capital. We see a gradual transition of aggregate capital towards a higher level and a rightwards

[30]The plots shows that the first-order approximation under our method and the approximation of Auclert et al. (2021) overlap. This overlap is reassuring that issues related to numerical derivatives and choice of asset grid are not quantitatively large.

movement of the the distribution of wealth.

7.2.2 Welfare from stabilization policies

Second-order approximations can be used to evaluate welfare effects of alternative stabilization policies, for instance, fiscal or monetary rules that describe how taxes or interest rates vary over business cycles. Here, we extend the baseline Krusell and Smith model to include a fiscal rule in form of a time varying labor tax

$$\tau_t = \tau_\Theta \Theta_t,$$

where τ_Θ is a scalar stabilization parameter. After-tax labor income is $(1-\tau_t)W_t \exp(\theta_{i,t})$ and tevenues from this tax are returned lump-sum T_t to the households.

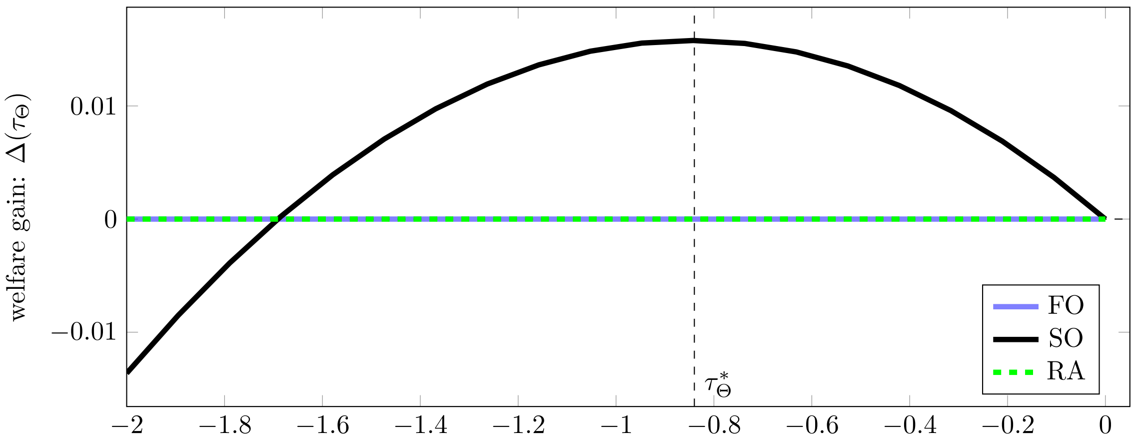
The magnitude of τ_Θ controls the transfer of resources across productive and unproductive households in response to aggregate shocks. For a given τ_Θ, define ergodic utilitarian welfare as $\mathbb{E}\mathcal{W}(\Theta, A, \Omega; \tau_\Theta) = \mathbb{E}\left(\int v(a, \theta, \Theta, A, \Omega; \tau_\Theta)\, d\Omega\right)$ where v is the value of an individual who starts with idiosyncratic states (a, θ) when the aggregate state is (Θ, A, Ω) under policy indexed by τ_Θ. To measure the welfare changes from a tax reform $\tau_\Theta > 0$ in interpretable units, we compute a scalar $\Delta(\tau_\Theta)$ which is the the common per-period percentage change in households' consumption relative to the allocation under the laissez-faire policy $\tau_\Theta = 0$ so that the ergodic welfare equals post-reform value $\mathbb{E}\mathcal{W}(\Theta, A, \Omega; \tau_\Theta)$. For preferences used in the baseline, this welfare-equivalent consumption change is given by $\Delta(\tau_\Theta) = \left[\frac{\mathbb{E}\mathcal{W}(\Theta, A, \Omega; \tau_\Theta)}{\mathbb{E}\mathcal{W}(\Theta, A, \Omega; \tau_\Theta = 0)}\right]^{\frac{1}{1-\gamma}} - 1$.

We follow steps from Section 3.4 and equation (17) to approximate $\mathbb{E}\mathcal{W}(\Theta, A, \Omega; \tau_\Theta)$ across different choices of τ_Θ. By extending x and X to include v and $\mathcal{W}$, respectively, and adding the Bellman equation that solves the value function v to the mapping F and the definition of welfare $\mathcal{W}$ to the mapping G, our framework computes welfare automatically. As mentioned before, it takes only a few seconds to calculate $\Delta(\tau_\Theta)$ for a given τ_Θ.

In Figure 3, we plot $\Delta(\tau_\Theta)$ as a function of the tax parameter τ_Θ. We see that the welfare gain is zero in the representative agent economy (Ricardian equivlance) and also to the first-order of approximation in the HA economy (certainty equivlance). At the second-order HA (black line), we see a meaningful welfare tradeoff across different values of τ_Θ. Making the tax policy more countercyclical initially raises welfare with a distinct maximum (denoted by τ_Θ^*) achieved at $\tau_\Theta = -.84$ which amounts to raising taxes by 0.84 percentage points for every percentage point decrease in TFP.

This application also serves as a valuable tool for illustrating the shortcomings associated with employing the histogram technique. In Section 6, we emphasized the consequences of naively extending the histogram approach, which overlook specific second-order terms. These second-order terms become particularly crucial when calculating welfare derived from stabilization policy, which is inherently a

Figure 3: WELFARE FROM STABILIZATION POLICY



Notes: The figure plots the $\Delta(\tau_\Theta)$, which is the welfare-equivalent consumption change relative to laissez-faire policy, $\tau_\Theta = 0$. The solid black line "SO" uses the second-order approximation from Section 3.4. The solid blue line "FO" uses the first-order approximation from Section 3.3 and the dashed green "RA" line plots the welfare gain using the representative agent counterpart targeted to the same set of aggregate moments.

second-order object. We find that the optimal cyclicality parameter using the histogram method to compute the welfare equals -1.04 which is substantially different from what we found with our method. It turns out that both, the magnitude of welfare corresponding to a particular τ_Θ and the gradient of welfare in relation to τ_Θ are inaccurate when utilizing the histogram approach.

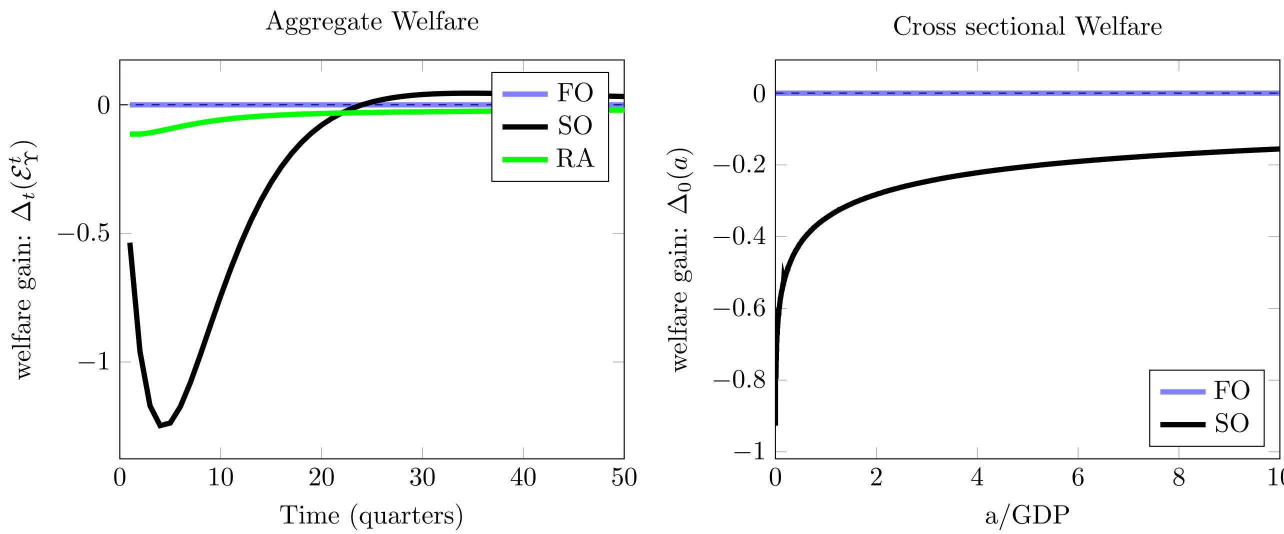
7.2.3 Stochastic Volatility

We next use techniques from Section 5.2 to study aggregate and distributional consequences of changes in macroeconomic risk. To do that, we first extend the baseline model to include equations (50)–(51) as the new process for aggregate shocks. This introduces two new parameters ρ_Υ and $\sigma^2_\Upsilon = var(\mathcal{E}_{\gamma,t})$. We use the fluctuations in CBOE Volatility Index (VIX) as a proxy for Υ_t and estimate the two parameters using quarterly data on VIX for the sample period 1990-2023. See Appendix E.3 for more details of the estimation.

Consider a path for the innovations to aggregate uncertainty $\tilde{\mathcal{E}}^t_\Upsilon$. Analogous to the previous section, we measure the effect of $\tilde{\mathcal{E}}^t_\Upsilon$ on aggregate welfare using $\Delta_t\left(\tilde{\mathcal{E}}^t_\Upsilon\right) = \left[\frac{\mathbb{E}_{\mathcal{E}^t}\mathcal{W}_t\left(\tilde{\mathcal{E}}^t_\Upsilon,\mathcal{E}^t\right)}{\mathbb{E}_{\left(\mathcal{E}^t,\mathcal{E}^t_\Upsilon\right)}\mathcal{W}_t\left(\mathcal{E}^t_\Upsilon,\mathcal{E}^t\right)}\right]^{\frac{1}{1-\gamma}} - 1$, which is the date t onwards welfare-equivalent change in household consumption. Since we have TFP shocks in the background, we integrate over the paths of $\mathcal{E}^t$.

In our sample, the VIX (see Figure 7 in the appendix) is fairly stable but features large spikes (about 5X increases) in 2008 after the Lehman collapse and in 2020 after the Covid pandemic. We are interested in studying the welfare effects—aggregate and distributional—of such spikes in uncertainty. To do that, we simulate a path for $\Delta_t\left(\tilde{\mathcal{E}}^t_\Upsilon\right)$ for the sequence $\tilde{\mathcal{E}}^t_\Upsilon = \left(\tilde{\mathcal{E}}_{\Upsilon,0}, 0, 0, 0, \ldots\right)$ where $\tilde{\mathcal{E}}_{\Upsilon,0}$ is

Figure 4: WELFARE EFFECTS OF UNCERTAINTY



Notes: The figure plots the aggregate and distributional welfare effects of an aggregate uncertainty corresponding to a one-time 5X increase in standard deviation of TFP. The left panel plots $\Delta_t\left(\tilde{\mathcal{E}}^t_\Upsilon\right)$, which is the aggregate welfare-equivalent consumption change after the spike in uncertainty. The right panel plots individual-level welfare-equivalent consumption change for households with asset-level a, $\int \Delta_0\left(a, \theta; \mathcal{E}^t_\Upsilon\right) d\Omega(\theta|a)$ to avoid the spike in uncertainty at $t = 0$. Individual assets (on the x-axis) are normalized by per-capita GDP.

chosen to match a 5X increase in the standard deviation of TFP and then the process Υ_t mean reverts with the estimated persistence ρ_Υ. In left panel of Figure 4, we observe that the shock leads to a decrease in aggregate welfare on impact by 0.53%. The effects are substantially amplified relative to the representative agent counterpart.

In addition to the impact on aggregate variables, our method allows us to investigate the effect of the shock on individual welfare. For each (a, θ), we compute $\Delta_0\left(a, \theta; \tilde{\mathcal{E}}^t_\Upsilon\right) = \left[\frac{\mathbb{E}_{\mathcal{E}^t} v_0\left(a,\theta;\mathcal{E}^t,\tilde{\mathcal{E}}^t_\Upsilon\right)}{\mathbb{E}_{\left(\mathcal{E}^t,\mathcal{E}^t_\Upsilon\right)} v_0\left(a,\theta;\mathcal{E}^t,\mathcal{E}^t_\Upsilon\right)}\right]^{\frac{1}{1-\gamma}} - 1$, which is per-period consumption change for a household with states (a, θ) at date $t = 0$ to avoid the spike in uncertainty. In right panel of Figure 4, we plot the welfare losses averaged across productivities for each level of assets/per capita GDP. The average welfare loss amounts to approximately half a percentage point of per-period consumption, and these losses range from 0.94% to 0.20% across the asset distribution. The most significant welfare losses are experienced by asset-poor agents who are closer to the borrowing constraints.

7.2.4 Portfolio choice

Finally, we illustrate the extension of our algorithm in Section 5.3 to capture portfolio choice. Extend the baseline Krusell and Smith model allowing agents to trade risk-free debt, b, which has a zero net supply, in addition to claims on risky capital whose market value we denote by k. Total wealth is

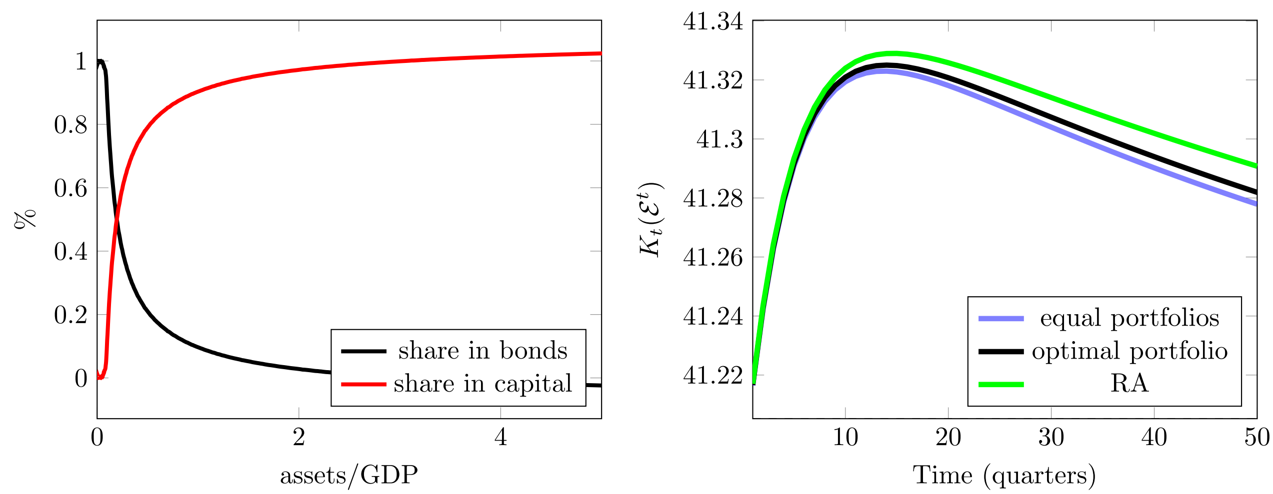
$a = k + b$. We impose a constraint that prevents households from short-selling capital.

The key computational step here is to construct $\left\{\mathsf{J}^{PP}_{t,s}\right\}_{t,s}$ and then apply Proposition 1^{PP} to compute the first-order responses to aggregates. This takes $0.78s$ more than than the time reported to compute the first-order responses without the portfolio choice. The first-order responses are then used to construct the zeroth-order portfolios using Lemma 7.

We now explore the predictions of the baseline Krusell and Smith model for the cross-sectional distribution of portfolios as well as the role of portfolios in shaping aggregate responses. In the left panel of Figure 5, we depict the distribution of household portfolios by assets normalized by per capita GDP. The model qualitatively aligns with the observed pattern (see Yogo and Wachter (2011) who use data from the Survey of Consumer Finances) wherein poorer households hold more bonds and wealthier households hold more stocks. Households closest to the borrowing constraint are most exposed to aggregate shocks, and they optimally reduce their exposure by adjusting their portfolios towards risk-free bonds.

Optimal portfolios matter even for a first-order approximation of aggregates. To see this, we simulate $K_t\left(\mathcal{E}^t\right)$ for the sequence $\mathcal{E}^t = (1, 0, 0, \ldots)$ and report the first-order approximation with optimal portfolio and compare it to the response if we force households to hold the same portfolios. In the right panel of Figure 5, we see that the responses with optimal portfolio are larger.

Figure 5: PORTFOLIOS



Notes: The left panel plots the cross distribution of portfolios by value of assets (normalized by per capita GDP). The right panel plots the first-order path of aggregate capital $K_t\left(\mathcal{E}^t\right)$ when households optimally chose the portfolio (black line) as in Section 5.3, the response of capital when households are forced to hold the same portfolio (blue line), and the response of capital under the representative agent economy (green line).

8 Conclusion

In this paper, we propose a novel perturbation technique to approximate a wide variety of stochastic heterogeneous-agent (HA) models. Our methods goes beyond the MIT shock approach prevalent in existing literature by employing higher-order approximations. Utilizing a directional derivative formulation, we demonstrate that all-order approximations can be represented using analytically derived coefficients that are straightforward to implement numerically. Our approach broadens the range of research questions that can be addressed within these model classes. We showcase the practicality of our method by applying it to examine welfare implications of stabilization policies, portfolio choice, and time-varying uncertainty in a calibrated economy.

References

Achdou, Yves, Jiequn Han, Jean Michel Lasry, Pierre Louis Lions, and Ben Moll. 2020. "Income and Wealth Distribution in Macroeconomics: a continuous-time approach." *Review of Economic Studies*.

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry, and Christian Wolf. 2018. "When Inequality Matters for Macro and Macro Matters for Inequality." *NBER Macroeconomics Annual*, 32(1): 1–75.

Alvarez, Fernando, and Francesco Lippi. 2022. "The Analytic Theory of a Monetary Shock." *Econometrica*, 90(4): 1653–1682.

Alvarez, Fernando, Francesco Lippi, and Panagiotis Souganidis. 2023. "Price Setting with Strategic Complementarities as a Mean Field Game."

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub. 2021. "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models." *Econometrica*, 89(5): 2375–2408.

Bayer, Christian, Benjamin Born, and Ralph Luetticke. 2022. "Shocks, Frictions, and Inequality in US Business Cycles." CEPR Discussion Papers 14364.

Benigno, Gianluca, Pierpaolo Benigno, and Salvatore Nistico. 2013. "Second-order approximation of dynamic models with time-varying risk." *Journal of Economic Dynamics and Control*, 37(7): 1231–1247.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J. Sargent. 2021. "Inequality, Business Cycles, and Monetary-Fiscal Policy." *Econometrica*, 89(6): 2559–2599.

Bigio, Saki, Galo Nuño, and Juan Passadore. 2023. "Debt-Maturity Management with Liquidity Costs."

Bilal, Adrien. 2023. "Solving Heterogeneous Agent Models with the Master Equation."

Blanchard, Olivier Jean, and Charles M Kahn. 1980. "The Solution of Linear Difference Models under Rational Expectations." *Econometrica* 1305–1311.

Boppart, Timo, Per Krusell, and Kurt Mitman. 2018. "Exploiting MIT Shocks in Heterogeneous-Agent Economies: The Impulse Response as a Numerical Derivative." *Journal of Economic Dynamics and Control*, 89 68–92.

Carroll, Christopher D. 2006. "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems." *Economics Letters*, 91(3): 312–320.

Childers, David. 2018. "On the Solution and Application of Rational Expectations Models with Function-Valued States."Technical report.

Childers, David, Jesús Fernández-Villaverde, Jesse Perla, Christopher Rackauckas, and Peifan Wu. 2022. "Differentiable State-Space Models and Hamiltonian Monte Carlo Estimation."Technical report, National Bureau of Economic Research.

Devereux, Michael B., and Alan Sutherland. 2011. "Country Portfolios in Open Economy Macro-Models." *Journal of the European Economic Association*, 9(2): 337–369.

Fernández-Villaverde, Jesús, Pablo Guerrón-Quintana, Juan F. Rubio-Ramírez, and Martin Uribe. 2011. "Risk Matters: The Real Effects of Volatility Shocks." *American Economic Review*, 101(6): 2530–61.

Fernández-Villaverde, Jesús, Juan Francisco Rubio-Ramirez, and Frank Schorfheide. 2016. "Solution and Estimation methods for DSGE models." 2 527–724.

Gornemann, Nils, Keith Kuester, and Makoto Nakajima. 2021. "Doves for the Rich, Hawks for the Poor? Distributional Consequences of Systematic Monetary Policy."

Han, Jiequn, Yucheng Yang et al. 2021. "Deepham: A global solution method for heterogeneous agent models with aggregate shocks." *arXiv preprint arXiv:2112.14377*.

Holmes, Mark H. 2012. *Introduction to perturbation methods.* 20: Springer Science & Business Media.

Jin, He-Hui, and Kenneth L Judd. 2002. "Perturbation Methods for General Dynamic Stochastic Models."Technical report.

Judd, Kenneth L. 1998. *Numerical Methods in Economics.* MIT Press Books: The MIT Press.

Kahou, Mahdi Ebrahimi, Jesús Fernández-Villaverde, Jesse Perla, and Arnav Sood. 2021. "Exploiting symmetry in high-dimensional dynamic programming."Technical report, National Bureau of Economic Research.

Kanwal, Ram P. 1998. *Generalized Functions: Theory and Technique.*: Springer Science & Business Media.

Kaplan, Greg, Benjamin Moll, and Giovanni L Violante. 2018. "Monetary Policy according to HANK." *American Economic Review*, 108(3): 697–743.

Kim, Jinill, Sunghyun Kim, Ernst Schaumburg, and Christopher A Sims. 2008. "Calculating and using Second-Order Accurate Solutions of discrete-time Dynamic Equilibrium Models." *Journal of Economic Dynamics and Control*, 32(11): 3397–3414.

Krusell, Per, and Anthony A Smith, Jr. 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 106(5): 867–896.

Lombardo, Giovanni. 2010. "On approximating DSGE models by series expansions."

Lombardo, Giovanni, and Harald Uhlig. 2018. "A Theory Of Pruning." *International Economic Review*, 59(4): 1825–1836.

Maliar, Lilia, Serguei Maliar, and Pablo Winant. 2021. "Deep learning for solving dynamic economic models.." *Journal of Monetary Economics*, 122 76–101.

Mertens, Thomas M., and Kenneth L. Judd. 2018. "Solving an Incomplete Markets Model with a Large Cross-section of Agents." *Journal of Economic Dynamics and Control*, 91 349–368, Special Issue in Honour of Prof. Carl Chiarella.

Reiter, Michael. 2009. "Solving Heterogeneous-Agent Models by Projection and Perturbation." *Journal of Economic Dynamics and Control*, 33(3): 649–665.

Reiter, Michael. 2023. "State Reduction and Second-Order Perturbations of Heterogeneous Agent Models."Technical report.

Schmitt-Grohé, Stephanie, and Martın Uribe. 2004. "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function." *Journal of Economic Dynamics and Control*, 28(4): 755–775.

Uhlig, Harald. 2001. "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily." In *Computational Methods for the Study of Dynamic Economies.*: Oxford University Press.

Viceira, Luis M. 2001. "Optimal Portfolio Choice for Long Horizon Investors with Nontradable Labor Income." *Journal of Finance*, 56(2): 433–470.

Winberry, Thomas. 2018. "A method for solving and estimating heterogeneous agent macro models." *Journal of Economic Dynamics and Control*, 89 68–92.

Yogo, Motohiro, and Jessica A. Wachter. 2011. "Why Do Household Portfolio Shares Rise in Wealth?" *Review of Financial Studies*, 24(11): 3929–3965.

Young, Eric R. 2010. "Solving the Incomplete Markets Model with Aggregate Uncertainty using the Krusell–Smith Algorithm and Non-Stochastic Simulations." *Journal of Economic Dynamics and Control*, 34(1): 36–41.

Online Appendix

A Section 2 Proofs

A.1 Proof of Lemma 1

Taking a first-order derivative of (8) and (9) around the $\sigma = 0$ steady state yields $\overline{Z}_{0,\sigma}(\mathcal{E}^0) = \hat{Z}_0\mathcal{E}_0$ and, for $t \geq 0$,

$$\overline{Z}_{t+1,\sigma}\left(\mathcal{E}^{t+1}\right) = \overline{Z}_Z \cdot \overline{Z}_{t,\sigma}\left(\mathcal{E}^{t-1}\right) + \hat{Z}_0\mathcal{E}_t + \overline{Z}_\sigma \tag{67}$$

$$\overline{X}_{t,\sigma}\left(\mathcal{E}^{t}\right) = \overline{X}_Z \cdot \overline{Z}_{t,\sigma}\left(\mathcal{E}^{t}\right) + \overline{X}_\sigma, \tag{68}$$

with $\hat{Z}_0$ and $\overline{Z}_Z$ being defined in the main text and $\overline{Z}_\sigma := \left[0, \overline{K}_\sigma\right]$. Our first step is to show that $\overline{X}_\sigma$ and $\overline{Z}_\sigma$ are both 0 which we codify in the following claim

Claim 1. The first derivatives with respect to σ, $\overline{X}_\sigma$ is 0.

Proof. Differentiating the G mapping w.r.t. σ yields $\mathsf{G}_Y\overline{Y}_\sigma = 0$, where $\overline{Y} = \left[0, 0, \overline{X}_\sigma, \overline{X}_\sigma + \overline{X}_Z\overline{Z}_\sigma\right]$ and $\overline{Z}_\sigma = [0, \mathsf{P}\overline{X}_\sigma]$. This system of equations is homogeneous of degree 1 in $\left(\overline{X}_\sigma\right)$ and, therefore, is solved by setting all terms to zero. □

Next we show the following claim relating $\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)$ to the directions $\hat{Z}_t$

Claim 2. For all t, $\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right) = \sum_{s=0}^{t} \hat{Z}_{t-s}\mathcal{E}_s$.

Proof. We proceed via induction as $\overline{Z}_{0,\sigma}(\mathcal{E}^0) = \hat{Z}_0\mathcal{E}_0$ implies it holds for $t = 0$. Assuming it holds for $t-1$ we have

$$\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right) = \overline{Z}_Z \cdot \left(\sum_{s=0}^{t-1} \hat{Z}_{t-1-s}\mathcal{E}_s\right) + \hat{Z}_0\mathcal{E}_0 = \left(\sum_{s=0}^{t-1} \hat{Z}_{t-s}\mathcal{E}_s\right) + \hat{Z}_0\mathcal{E}_0 = \sum_{s=0}^{t} \hat{Z}_{t-s}\mathcal{E}_s,$$

where in the second equality we used $\hat{Z}_{k+1} \equiv \overline{Z}_Z \cdot \hat{Z}_k$. □

Finally, substituting for $\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)$ in (8) completes the proof as $\overline{X}_Z \cdot \hat{Z}_{t-s} = \hat{X}_{t-s}$.

A.2 Proof of Proposition 1

Begin by differentiating (5) in direction $\hat{Z}_t$ to find$\mathsf{G}_Y\overline{Y}_Z \cdot \hat{Z}_t = 0$. To obtain $\overline{Y}_Z \cdot \hat{Z}$ we differentiate $\overline{Y}(Z;\sigma)$, equation (6), in direction $\hat{Z} = \left[\hat{\Theta}, \hat{K}\right]$ to get $\overline{Y}_Z \cdot \hat{Z} = \left[\hat{\Theta}, \hat{K}, \overline{X}_Z \cdot \hat{Z}, \overline{X}_Z \cdot \overline{Z}_Z \cdot \hat{Z}\right]$. Using $\hat{Z}_t = \left[\rho_\Theta^t, \hat{K}_{t-1}\right] = \left[\rho_\Theta^t, \mathsf{P}\hat{X}_{t-1}\right]$[31] we have $\hat{Y}_t = \overline{Y}_Z \cdot \hat{Z}_t = \left[\rho_\Theta^t, \mathsf{P}\hat{X}_{t-1}, \hat{X}_t, \hat{X}_{t+1}\right]$.

Finally, differentiating $\overline{Z}_t$ defined in Assumption 1 implies that $\overline{Z}_{t,Z} \cdot \hat{Z}_0 = \hat{Z}_t$. Therefore, the first derivative of Assumption 1(b) implies $\lim_{t\to\infty} \hat{Z}_t = 0$, which in turn implies that $\lim_{t\to\infty} \hat{X}_t = \lim_{t\to\infty} \overline{X}_Z \cdot \hat{Z}_t = 0$.

[31] For $t > 0$ we have $\hat{K}_t = \mathsf{P}\overline{X}_Z \cdot \hat{Z}_{t-1} = \mathsf{P}\hat{X}_{t-1}$ while for $t = 0$ we have $\hat{K}_0 = 0$ by definition and impose the restriction $\mathsf{P}\hat{X}_{-1} := 0$

A.3 Proof of Lemma 2

We proceed by taking a second-order derivatives of (9) and (8) w.r.t. σ to find $\overline{Z}_{0,\sigma\sigma}\left(\mathcal{E}^0\right)=0$ and [32]

$$\overline{Z}_{t+1,\sigma\sigma}\left(\mathcal{E}^{t+1}\right)=\overline{Z}_Z\cdot\overline{Z}_{t,\sigma\sigma}\left(\mathcal{E}^t\right)+\overline{Z}_{ZZ}\cdot\left(\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right),\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)\right)+\overline{Z}_{\sigma\sigma} \tag{69}$$

$$\overline{X}_{t,\sigma\sigma}\left(\mathcal{E}^t\right)=\overline{X}_Z\cdot\overline{Z}_{t,\sigma\sigma}\left(\mathcal{E}^t\right)+\overline{X}_{ZZ}\cdot\left(\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right),\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)\right)+\overline{X}_{\sigma\sigma} \tag{70}$$

where $\overline{Z}_{ZZ}$ is defined in the main text and $\overline{Z}_{\sigma\sigma}=\left[0,\mathsf{P}\overline{X}_{\sigma\sigma}\right]^T$. We begin by showing the following claim relating $\overline{Z}_{t,\sigma\sigma}\left(\mathcal{E}^t\right)$ to the directions $\hat{Z}_{t,k}$ and $\hat{Z}_{\sigma\sigma,t}$.

Claim 3. For all t

$$\overline{Z}_{t,\sigma\sigma}\left(\mathcal{E}^t\right)=\hat{Z}_{\sigma\sigma,t}+\sum_{s=0}^{t}\sum_{m=0}^{t}\hat{Z}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m \tag{71}$$

Proof. We proceed by induction. As $\hat{Z}_{\sigma\sigma,0}=\hat{Z}_{0,0}=0$ we conclude that equation (8) holds for $t=0$ since $\overline{Z}_{0,\sigma\sigma}\left(\mathcal{E}^0\right)=0$. Assuming (71) holds for $t-1$ we have

$$\begin{aligned}\overline{Z}_{t,\sigma\sigma}\left(\mathcal{E}^t\right)&=\overline{Z}_Z\cdot\left(\hat{Z}_{\sigma\sigma,t-1}+\sum_{s=0}^{t-1}\sum_{m=0}^{t-1}\hat{Z}_{t-1-s,t-1-m}\mathcal{E}_s\mathcal{E}_m\right)+\overline{Z}_{ZZ}\cdot\left(\sum_{s=0}^{t-1}\hat{Z}_{t-1-s}\mathcal{E}_s,\sum_{m=0}^{t-1}\hat{Z}_{t-1-m}\mathcal{E}_m\right)+\overline{Z}_{\sigma\sigma}\\&=\overline{Z}_Z\cdot\hat{Z}_{\sigma\sigma,t-1}+\overline{Z}_{\sigma\sigma}+\sum_{s=0}^{t-1}\sum_{m=0}^{t-1}\left(\overline{Z}_Z\cdot\hat{Z}_{t-1-s,t-1-m}+\overline{Z}_{ZZ}\cdot\left(\hat{Z}_{t-1-s},\hat{Z}_{t-1-m}\right)\right)\mathcal{E}_s\mathcal{E}_m\\&=\hat{Z}_{\sigma\sigma,t}+\sum_{s=0}^{t}\sum_{m=0}^{t}\hat{Z}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m\end{aligned}$$

where in the second equality we used the fact that $\overline{Z}_{ZZ}$ is a bi-linear mapping and in the third equality we use the recursive definitions of $\hat{Z}_{\sigma\sigma,t}$ and $\hat{Z}_{t,k}$, and $\hat{Z}_{0,0}=0$. $\square$

Finally we plug in for $\overline{Z}_{t,\sigma\sigma}\left(\mathcal{E}^t\right)$ and $\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)$ in equation (70) to find

$$\begin{aligned}\overline{X}_{t,\sigma\sigma}\left(\mathcal{E}^t\right)&=\overline{X}_Z\cdot\left(\hat{Z}_{\sigma\sigma,t}+\sum_{s=0}^{t}\sum_{m=0}^{t}\hat{Z}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m\right)+\overline{X}_{ZZ}\cdot\left(\sum_{s=0}^{t}\hat{Z}_{t-s}\mathcal{E}_s,\sum_{m=0}^{t}\hat{Z}_{t-m}\mathcal{E}_m\right)+\overline{X}_{\sigma\sigma}\\&=\overline{X}_Z\cdot\hat{Z}_{\sigma\sigma,t}+\overline{X}_{\sigma\sigma}+\sum_{s=0}^{t}\sum_{m=0}^{t}\left(\overline{X}_Z\cdot\hat{Z}_{t-s,t-m}+\overline{X}_{ZZ}\cdot\left(\hat{Z}_{t-s},\hat{Z}_{t-m}\right)\right)\mathcal{E}_s\mathcal{E}_m\\&=\hat{X}_{\sigma\sigma,t}+\sum_{s=0}^{t}\sum_{m=0}^{t}\hat{X}_{ZZ,t-s,t-m}\mathcal{E}_s\mathcal{E}_m\end{aligned}$$

which completes the proof.

A.4 Proof of Proposition 2

Begin by differentiating equation (5) twice in directions $\hat{Z}_t$ and $\hat{Z}_{t+k}$ and add to it the derivative of (5) in direction $\hat{Z}_{t,t+k}$ to find $\mathsf{G}_Y\hat{Y}_{t,t+k}+\mathsf{G}_{YY}\cdot\left(\overline{Y}_Z\cdot\hat{Z}_t,\overline{Y}_Z\cdot\hat{Z}_{t+k}\right)=0$, where $\hat{Y}_{t,t+k}:=\overline{Y}_{ZZ}\cdot\left(\hat{Z}_t,\hat{Z}_{t+k}\right)+$

[32] There are also $\overline{X}_{\sigma Z}$ and $\overline{Z}_{\sigma Z}$ terms but they are 0 following the same logic as $\overline{X}_\sigma$ and $\overline{Z}_\sigma$ being 0 in the proof of Lemma 1.

$\overline{Y}_Z \cdot \hat{Z}_{t,t+k}$. Differentiating $\overline{Y}$ twice in direction $\hat{Z}_t$ and $\hat{Z}_{t+k}$ implies

$$\overline{Y}_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right) = \left[0, 0, \overline{X}_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right), \overline{X}_{ZZ} \cdot \left(\hat{Z}_{t+1}, \hat{Z}_{t+1+k}\right) + \overline{X}_Z \cdot \overline{Z}_Z \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right)\right]^T.$$

If we add to it $\overline{Y}_Z \cdot \hat{Z}_{t,t+k} = \left[0, \mathsf{P}\hat{X}_{t-1,t-1+k}, \overline{X}_Z \cdot \hat{Z}_{t,t+k}, \overline{X}_Z \cdot \overline{Z}_Z \cdot \hat{Z}_{t,t+k}\right]^T$ we find

$$\hat{Y}_{t,t+k} = \left[0, \mathsf{P}\hat{X}_{t-1,t-1+k}, \hat{X}_{t,t+k}, \hat{X}_{t+1,t+1+k}\right]$$

as desired. The same limiting arguments as Proposition 1 imply $\lim_{t\to 0} \hat{X}_{t,t+k} = 0$.

Next differentiating equation (5) twice with respect to σ and add to it the derivative of (5) in direction $\hat{Z}_{\sigma\sigma,t}$ to find$\mathsf{G}_Y\hat{Y}_{\sigma\sigma,t} = 0$ where $\hat{Y}_{\sigma\sigma,t} := \overline{Y}_{\sigma\sigma} + \overline{Y}_Z \cdot \hat{Z}_{\sigma\sigma,t}$. Differentiating $\overline{Y}$ twice with respect to σ yields

$$\overline{Y}_{\sigma\sigma} = \left[0, 0, \overline{X}_{\sigma\sigma}, \overline{X}_{\sigma\sigma} + \overline{X}_Z \cdot \overline{Z}_{\sigma\sigma} + \mathbb{E}_{\mathcal{E}}\left[\overline{X}_{ZZ} \cdot \left(\hat{Z}_0\mathcal{E}, \hat{Z}_0\mathcal{E}\right)\right]\right]^T.$$

Add to it

$$\overline{Y}_Z \cdot \hat{Z}_{\sigma\sigma,t} = \left[0, \mathsf{P}\hat{X}_{\sigma\sigma,t-1}, \overline{X}_Z \cdot \hat{Z}_{\sigma\sigma,t}, \overline{X}_Z \cdot \overline{Z}_Z \cdot \hat{Z}_{\sigma\sigma,t}\right]^T$$

to find $\hat{Y}_{\sigma\sigma,t} = \left[0, \mathsf{P}\hat{X}_{\sigma\sigma,t-1}, \hat{X}_{\sigma\sigma,t}, \hat{X}_{\sigma\sigma,t+1} + \hat{X}_{0,0}\text{var}(\mathcal{E})\right]$. Finally, as $\hat{Z}_{\sigma\sigma,t} - \hat{Z}_{\sigma\sigma,t-1} = \overline{Z}_Z \cdot \left(\hat{Z}_{\sigma\sigma,t} - \hat{Z}_{\sigma\sigma,t-1}\right)$ with $\hat{Z}_{\sigma\sigma,0} - \hat{Z}_{\sigma\sigma,-1} := \overline{Z}_{\sigma\sigma}$, we can conclude that Assumption 1(b) implies that $\lim_{t\to\infty} \hat{Z}_{\sigma\sigma,t} - \hat{Z}_{\sigma\sigma,t-1}$ and thus $\lim_{t\to\infty} \hat{X}_{\sigma\sigma,t} - \hat{X}_{\sigma\sigma,t-1} = 0$.

B Section 3 Proofs

B.1 Non-Negativity Constraints

We are interested in modeling the behavior of agents facing occasionally binding constraints. These constraints necessarily result in optimality conditions that must represented by both equality and inequality constraints. We will show here that for our approximation the inequality constraints are redundant.

We'll focus on the problem of one occasionally binding constraint, which extends automatically to multiple occasionally binding constraints. Without loss of generality, any occasionally binding constraint can be written as $k_{i,t} \geq 0$ by suitably defining the variable $k_{i,t}$. Letting $\xi_{i,t}$ be the multiplier on that constraint, the full set of optimality conditions can be written as the complimentary slackness condition $\zeta_{i,t}k_{i,t} = 0$ along with the non-negativity constraints $k_{i,t} \geq 0$ and $\zeta_{i,t} \geq 0$. In our recursive formulation, these can be expressed as

$$\overline{\zeta}(a, \theta, Z; \sigma)\overline{k}(a, \theta, Z; \sigma) = 0 \text{ and } \overline{k}(a, \theta, Z; \sigma) \geq 0 \text{ and } \overline{\zeta}(a, \theta, Z; \sigma) \geq 0. \tag{72}$$

These constraints in (72) are all satisfied in the deterministic steady state so that both $\overline{k}(a, \theta) \geq 0$ and $\overline{\zeta}(a, \theta) \geq 0$ for all a, θ. Moreover, the points of non-differentiability $\overline{\kappa}(\theta)$ are defined as the only points for which both $\overline{k}(a, \theta) = \overline{\zeta}(a, \theta) = 0$.

For any (a, θ) not on a kink, we can differentiate (72) in direction $\hat{Z}$. Differentiating the equality constraint always implies

$$\overline{\zeta}(a, \theta)\overline{k}_Z(a, \theta) \cdot \hat{Z} + \overline{\zeta}_Z(a, \theta) \cdot \hat{Z}\overline{k}(a, \theta) = 0. \tag{73}$$

As (a, θ) is not on a kink there are two possibilities: either $\overline{k}(a, \theta) > 0$ or $\overline{\zeta}(a, \theta) > 0$. Well assume $\overline{k}(a, \theta) > 0$ as the other case is symmetric. This implies that $\overline{\zeta}(a, \theta) = 0$, and thus (73) simplifies to $\overline{\zeta}_Z(a, \theta) \cdot \hat{Z} = 0$.

Turning now to the inequality constraints, the constraint $\overline{k}(a, \theta, Z; \sigma) \geq 0$ implies that, locally, any value of $\overline{k}_Z(a, \theta) \cdot \hat{Z}$ is valid since $\overline{k}(a, \theta) > 0$ implies that $\overline{k}(a, \theta, Z; \sigma) > 0$ for some neighborhood around $(Z^*, 0)$. However, as $\overline{\zeta}(a, \theta) = 0$, the third inequality constraint requires that $\overline{\zeta}_Z(a, \theta) \cdot \hat{Z} = 0$ as otherwise

$\overline{\zeta}(a,\theta,Z;\sigma) \geq 0$ would be violated by moving in an appropriate direction $\hat{Z}$. As $\overline{\zeta}_Z(a,\theta)\cdot\hat{Z} = 0$ is already guaranteed by the complimentary slackness condition, this implies that the additional inequality constraints are redundant.

B.2 Derivatives of Kinks and Generalized Functions

Assumption 2(a) states that the policy rules $(a,\theta,Z;\sigma)$ are smooth everywhere except for the locations $\overline{\kappa}_j(\theta,Z;\sigma)$. For the remainder of the appendix, we will assume a single kink but the results generalize directly for multiple kinks. All the additional terms induced by kinks are replaced by sums.

As the classical derivatives, e.g. $\mathring{\overline{x}}_{aa}(a,\theta)$, are not defined at those kinks, and for the purposes of integration, we represent them as generalized functions. We will find it convenient to use the notation $\overline{x}^{\Delta}(\theta) \equiv \lim_{a\downarrow\overline{\kappa}_j(\theta)}\overline{x}(a,\theta) - \lim_{a\uparrow\overline{\kappa}_j(\theta)}\overline{x}(a,\theta)$ to represent the size of the discontinuity at the kink. Fore conciseness, we'll define the upper and lower limits w.r.t. a as $\overline{x}^{+}(a,\theta) = \lim_{h\downarrow 0}\overline{x}(a+h,\theta)$ and $\overline{x}^{-}(a,\theta) = \lim_{h\uparrow 0}\overline{x}(a+h,\theta)$ respectively. Continuity of the policy rules implies that $\overline{x}^{\Delta}(\theta) = \overline{x}^{+}(\overline{\kappa}(\theta),\theta) - \overline{x}^{-}(\overline{\kappa}(\theta),\theta) = 0$, but the derivatives themselves are allowed to to be discontinuous at the kink: $\overline{x}_a^{\Delta}(\theta) \neq 0$.

Before formally studying the distributional derivatives, it is necessary to understand how the kinks themselves respond to the shocks. Continuity of the policy rules allows us to get the following relationship the derivative of the kink, $\overline{\kappa}_Z(\theta)\cdot\hat{Z}$, and the size of the discontinuity of the derivative of the policy rules at that kink, $\overline{x}_Z^{\Delta}(\theta)\cdot\hat{Z}$.

Claim 4. For all t, the derivatives of the kinks satisfy $\overline{x}_a^{\Delta}(\theta)\overline{\kappa}_Z(\theta)\cdot\hat{Z} = -\overline{x}_Z^{\Delta}(\theta)\cdot\hat{Z}$, and, in particular, $\overline{\kappa}_Z(\theta)\cdot\hat{Z} = -\overline{a}_a^{\Delta}(\theta)^{-1}\overline{a}_Z^{\Delta}(\theta)\cdot\hat{Z}$.

Proof. Continuity implies that $\overline{x}^{+}(\overline{\kappa}(\theta,Z),\theta,Z) = \overline{x}^{-}(\overline{\kappa}(\theta,Z),\theta,Z)$. Differentiating with respect to Z in direction $\hat{Z}_t$ at $\sigma = 0$ yields

$$\overline{x}_a^{+}(\overline{\kappa}(\theta),\theta)\overline{\kappa}_Z(\theta)\cdot\hat{Z} + \overline{x}_Z^{+}(\overline{\kappa}(\theta),\theta)\cdot\hat{Z} = \overline{x}_a^{-}(\overline{\kappa}(\theta),\theta)\overline{\kappa}_Z(\theta)\cdot\hat{Z} + \overline{x}_Z^{-}(\overline{\kappa}(\theta),\theta)\cdot\hat{Z},$$

which implies that $\overline{x}_a^{\Delta}(\theta)\overline{\kappa}_Z(\theta)\cdot\hat{Z} = -\overline{x}_Z^{\Delta}(\theta)\cdot\hat{Z}$. Applying p to both sides and dividing by $\overline{a}_a^{\Delta}(\theta)$ completes the proof. □

The distributional derivates themselves are defined by how the operate as linear functionals over a space of smooth test functions, φ, with compact support. We use these definitions to establish the following relationships

Claim 5. For all t,k distributional derivatives of $\tilde{x}$ satisfy

$$\begin{aligned}
\overline{x}_Z(a,\theta)\cdot\hat{Z} &= \mathring{\overline{x}}_Z(a,\theta)\cdot\hat{Z}\\
\overline{x}_a(a,\theta) &= \mathring{\overline{x}}_a(a,\theta)\\
\overline{x}_{aa}(a,\theta) &= \mathring{\overline{x}}_{aa}(a,\theta) + \overline{x}_a^{\Delta}(\theta)\delta(a-\overline{\kappa}(\theta))\\
\overline{x}_{aZ}(a,\theta)\cdot\hat{Z} &= \mathring{\overline{x}}_{aZ}(a,\theta)\cdot\hat{Z} + \overline{x}_Z^{\Delta}(\theta)\cdot\hat{Z}\delta(a-\overline{\kappa}(\theta))\\
\overline{x}_{ZZ}(a,\theta)\cdot\left(\hat{Z}',\hat{Z}''\right) &= \mathring{\overline{x}}_{ZZ}(a,\theta)\cdot\left(\hat{Z}',\hat{Z}''\right) + \overline{x}_a^{\Delta}(\theta)\overline{\kappa}_Z(\theta)\cdot\hat{Z}'\overline{\kappa}_Z(\theta)\cdot\hat{Z}''\delta(a-\overline{\kappa}(\theta))
\end{aligned}$$

Proof. The distribution derivative $\overline{x}_a(a,\theta)$ is defined by[33]

$$\iint \overline{x}_a(a,\theta)\varphi(a,\theta)da d\theta = -\iint \overline{x}(a,\theta)\varphi_a(a,\theta)da d\theta = -\int\int_{-\infty}^{\overline{\kappa}(\theta)}\overline{x}(a,\theta)\varphi_a(a,\theta)da d\theta + \int\int_{\overline{\kappa}(\theta)}^{\infty}\overline{x}(a,\theta)\varphi_a(a,\theta)da d\theta$$

[33] To concisely represent these integrals we use the convention that $\overline{a}^{\vee,0}(\theta) = -\infty$ and $\overline{a}^{\vee,N^{\vee}+1}(\theta) = \infty$

for any test function φ. On each of these intervals the functions are smooth so we can apply integration by parts to get

$$\iint \overline{x}_a(a,\theta)\varphi(a,\theta)dad\theta = \iint \mathring{\overline{x}}_a(a,\theta)\varphi(a,\theta)dad\theta + \int \overline{x}^{\Delta}(\theta)\varphi(\overline{\kappa}(\theta),\theta)d\theta = \iint \mathring{\overline{x}}_a(a,\theta)\phi(a,\theta)dad\theta$$

where the last equality used continuity. This implies $\overline{x}_a(a,\theta) = \mathring{\overline{x}}_a(a,\theta)$.

Next we turn to $\overline{x}_{aa}(a,\theta)$, which is defined by

$$\iint \overline{x}_{aa}(a,\theta)\varphi(a,\theta)dad\theta = -\iint \overline{x}_a(a,\theta)\varphi_a(a,\theta)dad\theta = -\iint \mathring{\overline{x}}_a(a,\theta)\varphi_a(a,\theta)dad\theta$$

Splitting up the integral over a we have

$$\begin{aligned}\iint \overline{x}_{aa}(a,\theta)\varphi(a,\theta)dad\theta &= -\int\int_{-\infty}^{\overline{\kappa}(\theta)} \mathring{\overline{x}}_a(a,\theta)\varphi_a(a,\theta)dad\theta - \int\int_{\overline{\kappa}(\theta)}^{\infty} \mathring{\overline{x}}_a(a,\theta)\varphi_a(a,\theta)dad\theta \\ &= \iint \mathring{\overline{x}}_{aa}(a,\theta)\varphi(a,\theta)dad\theta + \int \overline{x}_a^{\Delta}(\theta)\varphi(\overline{\kappa}(\theta),\theta)d\theta,\end{aligned}$$

which implies $\overline{x}_{aa}(a,\theta) = \mathring{\overline{x}}_{aa}(a,\theta) + \overline{x}_a^{\Delta}(\theta)\delta(a-\overline{\kappa}(\theta))$.

Next we define the distributional derivative $\overline{x}_Z(a,\theta)\cdot\hat{Z}$ by

$$\int \overline{x}_Z(a,\theta)\cdot\hat{Z}\varphi(a,\theta)dad\theta = \left(\iint \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta\right)_Z\cdot\hat{Z}.$$

As

$$\iint \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta = \int\int_{-\infty}^{\overline{\kappa}(\theta,Z)} \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta + \int\int_{\overline{\kappa}(\theta,Z)}^{\infty} \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta$$

when we take the derivative we get

$$\left(\iint \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta\right)_Z\cdot\hat{Z} = \iint \mathring{\overline{x}}_Z(a,\theta)\cdot\hat{Z}\varphi(a,\theta)dad\theta - \int \overline{x}^{\Delta}(\theta)\overline{\kappa}_Z(\theta)\cdot\hat{Z}\varphi(\overline{\kappa}(\theta),\theta)d\theta$$

which implies that $\overline{x}_Z(a,\theta)\cdot\hat{Z} = \mathring{\overline{x}}_Z(a,\theta)$ as $\overline{x}^{\Delta}(\theta) = 0$.

The distributional derivative $\overline{x}_{aZ}(a,\theta)\cdot\hat{Z}$ is defined by

$$\iint \overline{x}_{aZ}(a,\theta)\cdot\hat{Z}\varphi(a,\theta)dad\theta = -\iint \overline{x}_Z(a,\theta)\cdot\hat{Z}\varphi_a(a,\theta)dad\theta$$

for any test function φ. As $\overline{x}_Z(a,\theta)\cdot\hat{Z} = \mathring{\overline{x}}_Z(a,\theta)\cdot\hat{Z}$ we have

$$\begin{aligned}-\iint \overline{x}_Z(a,\theta)\cdot\hat{Z}\varphi_a(a,\theta)dad\theta &= -\int\int_{-\infty}^{\overline{\kappa}(\theta)} \mathring{\overline{x}}_Z(a,\theta)\cdot\hat{Z}\varphi_a(a,\theta)dad\theta + \int\int_{\overline{\kappa}(\theta)}^{\infty} \mathring{\overline{x}}_Z(a,\theta)\cdot\hat{Z}\varphi_a(a,\theta)dad\theta \\ &= \iint \mathring{\overline{x}}_{aZ}(a,\theta)\cdot\hat{Z}\varphi(a,\theta)dad\theta + \int \overline{x}_Z^{\Delta}(\theta)\cdot\hat{Z}\varphi(\overline{\kappa}(\theta),\theta)d\theta,\end{aligned}$$

which implies $\overline{x}_{aZ}(a,\theta)\cdot\hat{Z} = \mathring{\overline{x}}_{aZ}(a,\theta)\cdot\hat{Z} + \overline{x}_Z^{\Delta}(\theta)\cdot\hat{Z}\delta(a-\overline{\kappa}(\theta))$.

Finally the distributional derivative, $\overline{x}_{ZZ,}(a,\theta)\cdot\left(\hat{Z}',\hat{Z}''\right)$, is defined by

$$\int \overline{x}_{ZZ,t,k}(a,\theta)\cdot\left(\hat{Z}',\hat{Z}''\right)\varphi(a,\theta)dad\theta = \left(\iint \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta\right)_{ZZ}\cdot\left(\hat{Z}',\hat{Z}''\right)$$

As

$$\left(\iint \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta\right)_{ZZ}\cdot\left(\hat{Z}',\hat{Z}''\right)=\left(\int\int_{-\infty}^{\overline{\kappa}(\theta,Z)}\mathring{\overline{x}}_Z(a,\theta,Z)\cdot\hat{Z}'\varphi(a,\theta)dad\theta\right)_Z\cdot\hat{Z}''$$
$$+\left(\int\int_{\overline{\kappa}(\theta,Z)}^{\infty}\mathring{\overline{x}}_Z(a,\theta,Z)\cdot\hat{Z}'\varphi(a,\theta)dad\theta\right)_Z\cdot\hat{Z}''$$
$$=\iint\mathring{\overline{x}}_{ZZ}(a,\theta)\cdot\left(\hat{Z}',\hat{Z}''\right)\varphi(a,\theta)dad\theta-\int\overline{x}_Z^{\Delta}(\theta)\cdot\hat{Z}'\overline{\kappa}_Z(\theta)\cdot\hat{Z}''\varphi\left(\overline{\kappa}(\theta),\theta\right)d\theta$$

which implies $\overline{x}_{ZZ}(a,\theta)\cdot\left(\hat{Z}',\hat{Z}''\right)=\mathring{\overline{x}}_{ZZ}(a,\theta)\cdot\left(\hat{Z}',\hat{Z}''\right)+\overline{x}_a^{\Delta}(\theta)\overline{\kappa}_Z(\theta)\cdot\hat{Z}'\overline{\kappa}_Z(\theta)\cdot\hat{Z}''\delta(a-\overline{\kappa}(\theta))$ as $\overline{x}_Z^{\Delta}(\theta)\cdot\hat{Z}'=-\overline{x}_a^{\Delta}(\theta)\overline{\kappa}_Z(\theta)\cdot\hat{Z}'$ from Claim 4. □

The distributional derivatives in Claim 5 provide a succinct way to summarize how changes in the location of the kink affect derivatives of integrals over individual policies. To keep the the analysis in this appendix as accessible as possible we'll derive all our main results without explicitly the generalized derivatives of $\overline{x}$. Instead, we will explicitly track the limits of integration and only summarize our results at the end using these δ-functions. As the distributional and classical derivatives align to first order we will use $\mathring{\overline{x}}_Z\cdot\hat{Z}$ and $\overline{x}_Z\cdot\hat{Z}$ interchangeably. We will only explicitly emphasize the classical derivative at second order.

Finally, we want to highlight an important feature of these additional terms that arise from kinked policy functions. Namely, they can always be determined from lower order derivatives. We see this in all of the generalized second derivatives, who's δ-function components depend only on first derivatives. This implies that all of the δ function components in the second order derivatives can be determined before the classical second order derivatives, are found.

B.3 Derivations of Equation (30)

For the direction $\hat{Z}_t=[\rho_\Theta^t,\mathsf{P}\hat{X}_{t-1},\hat{\Omega}_t]^\intercal$, determining $\left(\int\overline{x}d\Omega\right)_Z\cdot\hat{Z}_t$ is equivalent to differentiating $\int\overline{x}\left(a,\theta,Z^*+\alpha\hat{Z}_t\right)d\left(\Omega^*+\alpha\hat{\Omega}_t\right)$ which gives

$$\left(\int\overline{x}d\Omega\right)_Z\cdot\hat{Z}_t=\left(\int\overline{x}d\Omega^*\right)_Z\cdot\hat{Z}_t+\int\overline{x}d\hat{\Omega}_t=\int\hat{x}_td\Omega^*+\int\overline{x}d\hat{\Omega}_t.$$

B.4 Proof of Lemma 3

We begin by differentiating the F mapping, equation (25), in direction $\hat{Z}_t$ at a point not on the kinks. Doing so yields

$$\mathsf{F}_x(a,\theta)\overline{x}_Z(a,\theta)\cdot\hat{Z}_t+\mathsf{F}_Y(a,\theta)\overline{Y}_Z\cdot\hat{Z}_t+\mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right]\right)_Z\cdot\hat{Z}_t=0$$

where $\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right]=\int\overline{x}\left(\overline{a}(a,\theta,Z),\rho_\theta\theta+\varepsilon,\overline{Z}(Z)\right)\mu(\epsilon)d\epsilon$. Applying the derivative yields

$$\left(\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right]\right)_Z\cdot\hat{Z}_t=\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{p}\overline{x}_Z(a,\theta)\cdot\hat{Z}_t+\mathbb{E}_\varepsilon\left[\overline{x}_Z\cdot\overline{Z}_Z\cdot\hat{Z}_t|a,\theta\right]$$

Replacing $\overline{x}_Z\cdot\hat{Z}_t=\hat{x}_t$, $\overline{Y}_Z\cdot\hat{Z}_t=\hat{Y}_t$ and $\hat{Z}_{t+1}=\overline{Z}_Z\cdot\hat{Z}_t$ we get the difference equation

$$\mathsf{F}_x(a,\theta)\hat{x}_t(a,\theta)+\mathsf{F}_Y(a,\theta)\hat{Y}_t+\mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{p}\hat{x}_t(a,\theta)+\mathbb{E}_\varepsilon\left[\hat{x}_{t+1}|a,\theta\right]\right)=0.\tag{74}$$

Our claim is that $\hat{x}_t = \sum_{s=0}^{\infty} \mathsf{x}_s \hat{Y}_{t+s}$ solves this equation where x_s are defined via (32) and (33). To see this, note that

$$\begin{aligned}\mathsf{F}_{x^e}(a,\theta)\mathbb{E}_{\varepsilon}\left[\hat{x}_{t+1}|a,\theta\right] &= \sum_{s=0}^{\infty} \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_{\varepsilon}\left[\mathsf{x}_s|a,\theta\right]\hat{Y}_{t+1+s} \\ &= -\left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_{\varepsilon}\left[\overline{x}_a|a,\theta\right]\mathsf{P}\right)\sum_{s=1}^{\infty}\mathsf{x}_s(a,\theta)\hat{Y}_{t+s}\end{aligned}$$

where the second line comes from applying equation (33). Combined with equation (32) we have

$$\mathsf{F}_Y(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_{\varepsilon}\left[\hat{x}_{t+1}|a,\theta\right] = -\left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_{\varepsilon}\left[\overline{x}_a|a,\theta\right]\mathsf{P}\right)\hat{x}_t(a,\theta)$$

which guarantees (74) and completes the proof.

B.5 Proof of Lemma 4^{HA}

Differentiating the LoM, equation 27, in direction $\hat{Z}_t$ is equivalent to differentiating

$$\overline{\Omega}(Z^* + \alpha\hat{Z}_t)\langle a',\theta'\rangle = \int\int \iota\left(\bar{a}(a,\theta,Z^*+\alpha\hat{Z}_t) \le a'\right)\iota\left(\rho_\theta\theta + \epsilon \le \theta'\right)\mu(\epsilon)d\epsilon d\left(\Omega^* + \alpha\hat{\Omega}_t\right)\langle a,\theta\rangle$$

with respect to α. This yields

$$\begin{aligned}\overline{\Omega}_Z \cdot \hat{Z}_t\langle a',\theta'\rangle = &-\iint \delta\left(\overline{a}(a,\theta) - a'\right)\iota\left(\rho_\theta\theta + \epsilon \le \theta'\right)\mu(\epsilon)d\epsilon\hat{a}_t(a,\theta)d\Omega^*\langle a,\theta\rangle \\ &+ \iint \iota\left(\bar{a}(a,\theta) \le a'\right)\iota\left(\rho_\theta\theta + \epsilon \le \theta'\right)\mu(\epsilon)d\epsilon d\hat{\Omega}_t\langle a,\theta\rangle,\end{aligned}$$

where $\hat{a}_t = \mathsf{p}\hat{x}_t$. Applying $\frac{d}{d\theta'}$ to both sides yields

$$\begin{aligned}\frac{d}{d\theta'}\hat{\Omega}_{t+1}\langle a',\theta'\rangle &= -\int \overbrace{\delta\left(\overline{a}(a,\theta) - a'\right)\mu(\theta' - \rho_\theta\theta)}^{\overline{\Lambda}(a',\theta',a,\theta)}\hat{a}_t(a,\theta)d\Omega^*\langle a,\theta\rangle + \int \iota\left(\bar{a}(a,\theta) \le a'\right)\mu(\theta' - \rho_\theta\theta)d\hat{\Omega}_t\langle a,\theta\rangle \\ &= -\left(\mathcal{M}\cdot\overline{a}_{Z,t}\right)\langle a',\theta'\rangle + \int \overbrace{\delta\left(\overline{a}(a,\theta) - a'\right)\mu(\theta' - \rho_\theta\theta)}^{\overline{\Lambda}(a',\theta',a,\theta)}\bar{a}_a(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle dad\theta \\ &= -\left(\mathcal{M}\cdot\hat{a}_t\right)\langle a',\theta'\rangle + \left(\mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_t\right)\langle a',\theta'\rangle\end{aligned}$$

Where the second equality is achieved via integration by parts. To conclude, we will show that like ω^*, $\frac{d}{d\theta}\hat{\Omega}_t$ is also a generalized function of the form

$$\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle = \mathring{\omega}_t(a,\theta) + \hat{\xi}_{n,t}\delta(a - a_n^*). \tag{75}$$

We show this via induction

Claim 6. If y is a piecewise smooth with kinks at $\overline{\kappa}(\theta)$ then $\mathcal{M}\cdot\mathsf{y}$ is a generalized function with a finite number of mass points a_n^*.

Proof. From our definition of $\mathcal{M}$

$$\left(\mathcal{M}\cdot\mathsf{y}\right)\langle a',\theta'\rangle = \iint \overline{\Lambda}(a',\theta',a,\theta)\mathsf{y}(a,\theta)\omega^*(a,\theta)dad\theta.$$

As $\mathsf{y}(a,\theta)$ is only discontinuous at $\overline{\kappa}(\theta)$ and $\{(\overline{\kappa}(\theta),\theta)\}$ is measure 0 under ω^* guarantees that this integral is well-defined. Compare to ω^* which satisfies $\omega^*(a',\theta') = \iint \overline{\Lambda}(a',\theta',a,\theta)\mathsf{y}(a,\theta)\omega^*(a,\theta)dad\theta$. As $\mathsf{y}(a,\theta)$ is bounded over the support of ω^* we conclude that $\mathcal{M}\cdot\mathsf{y}$ is absolutely continuous with respect to ω^* and thus will only have a finite number of mass points at a_n^* . □

This claim implies that $\mathcal{M}\cdot\hat{a}_t$ is a generalized function with a finite number of mass-points at a_n^*. As $\hat{\Omega}_0 = 0$ we conclude that $\frac{d}{d\theta}\hat{\Omega}_1 = -\mathcal{M}\cdot\hat{a}_0$ is a generalized function with a finite number of mass-points at a_n^*. Our next claim allows us to extend this to all $\frac{d}{d\theta}\hat{\Omega}_t$ via induction. For the remainder of the proof we will use that $\frac{d}{d\theta}\hat{\Omega}_t$ has the form in (75)

Claim 7. If $\frac{d}{d\theta}\hat{\Omega}$ is a generalized function with a finite number of mass-points at a_n^*, then $\mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}$ is a generalized function with a finite number of mass-points at a_n^*.

Proof. Repeat the steps of the Claim 6 replacing y with $\overline{a}_a$ and $\mathring{\omega}^*(a,\theta) + \sum_n \xi_n^*(\theta)\delta(a-a_n^*)$ with $\frac{d}{d\theta}\hat{\Omega}$. □

B.6 Proof of Corollary 1

We start with our first claim

Claim 8. $\frac{d}{d\theta}\hat{\Omega}_t$ is given by $\frac{d}{d\theta}\hat{\Omega}_t = -\sum_{s=0}\mathsf{A}_{t,s}\hat{Y}_s$ where $\mathsf{A}_{t,s}$ is defined recursively by $\mathsf{A}_{0,s}=0$ and $\mathsf{A}_{t+1,s} = \mathcal{L}^{(a)}\cdot\mathsf{A}_{t,s} - \mathcal{M}\cdot\mathsf{px}_{s-t}$.

Proof. We proceed by induction. It's trivially true from $t=0$ as $\mathsf{A}_{0,s}=0$ and $\frac{d}{d\theta}\hat{\Omega}_0$. We then proceed by induction by substituting for $\hat{a}_t$

$$\frac{d}{d\theta}\hat{\Omega}_{t+1} = \mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_t - \sum_{j=0}^{\infty}\left(\mathcal{M}\cdot\mathsf{px}_j\right)\hat{Y}_{t+j} = \mathcal{L}^{(a)}\cdot\left(-\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\hat{Y}_s\right) - \sum_{s=0}^{\infty}\left(\mathcal{M}\cdot\mathsf{px}_{s-t}\right)\hat{Y}_s$$
$$= \sum_{s=0}^{\infty} -\left(\mathcal{L}^{(a)}\cdot\mathsf{A}_{t,s} + \mathcal{M}\cdot\mathsf{px}_{s-t}\right)\hat{Y}_s \equiv \sum_{s=0}^{\infty}\mathsf{A}_{t+1,s}\hat{Y}_s$$

where the second equality is achieved by letting $s=t+j$ and WLOG setting $\mathsf{x}_k=0$ for $k<0$. □

Applying integration by parts we have $\int\overline{x}d\hat{\Omega}_t = -\iint\overline{x}_a\frac{d}{d\theta}\hat{\Omega}_t dad\theta := -\mathcal{I}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_t$. From the proof of Lemma 4^{HA} we know that $\frac{d}{d\theta}\hat{\Omega}_t$ is a density with mass points at a finite number of points a_n^*, which implies that the set of points where $\overline{x}_a$ is not defined is measure zero under $\frac{d}{d\theta}\hat{\Omega}_t dad\theta$ so $\mathcal{I}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_t$ is well defined. and therefore $\int\overline{x}d\hat{\Omega}_t = \sum_{s=0}^{\infty}\left(\mathcal{I}^{(a)}\cdot\mathsf{A}_{t,s}\right)\hat{Y}_s$. We conclude by substituting for $\hat{x}_t$ and $\int\overline{x}d\hat{\Omega}_t$ in equation (30) to yield

$$\left(\int\overline{x}d\Omega\right)_Z\cdot\hat{Z}_t = \sum_{j=0}^{\infty}\int\mathsf{x}_j d\Omega^*\hat{Y}_{t+s} + \sum_{s=0}^{\infty}\left(\mathcal{I}^{(a)}\cdot\mathsf{A}_{t,s}\right)\hat{Y}_s = \sum_{s=0}^{\infty}\underbrace{\left(\int\mathsf{x}_{s-t}d\Omega^* + \mathcal{I}^{(a)}\cdot\mathsf{A}_{t,s}\right)}_{\mathsf{J}_{t,s}}\hat{Y}_s$$

as desired.

Finally, the recursive representation of $\mathsf{A}_{t,s}$ implies $\mathsf{A}_{t,s} = \sum_{k=1}^{t}\left(\mathcal{L}^{(a)}\right)^{t-k}\cdot\mathcal{M}\cdot\mathsf{px}_{s-k+1}$ and thus $\mathsf{A}_{t,s} - \mathsf{A}_{t-1,s-1} = \left(\mathcal{L}^{(a)}\right)^{t-1}\cdot\mathcal{M}\cdot\mathsf{px}_s$. This implies that $\mathsf{J}_{t,s}$ satisfies the desired recursion as $\mathsf{J}_{t,s} - \mathsf{J}_{t-1,s-1} = \mathcal{I}^{(a)}\cdot\left(\mathsf{A}_{t,s} - \mathsf{A}_{t-1,s-1}\right) = \mathcal{I}^{(a)}\cdot\left(\mathcal{L}^{(a)}\right)^{t-1}\cdot\mathcal{M}\cdot\mathsf{px}_s$.

B.7 Proof Of Proposition 1^{HA}

This is a direct result of combining Corollary 1 with equation (29).

B.8 Derivation of Equation (40)

Finding $\left(\int \overline{x} d\Omega\right)_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right)$ requires differentiating

$$\int \int_{-\infty}^{\overline{\kappa}(\theta,Z)} \overline{x}_Z(a,\theta,Z) \cdot \hat{Z}_t d\Omega\langle a,\theta\rangle + \int \int_{\overline{\kappa}(\theta,Z)}^{\infty} \overline{x}_Z(a,\theta,Z) \cdot \hat{Z}_t d\Omega\langle a,\theta\rangle$$
$$+ \int \int_{-\infty}^{\overline{\kappa}(\theta,Z)} \overline{x}(a,\theta,Z) d\hat{\Omega}_t\langle a,\theta\rangle + \int \int_{\overline{\kappa}(\theta,Z)}^{\infty} \overline{x}(a,\theta,Z) d\hat{\Omega}_t\langle a,\theta\rangle$$

in direction $\hat{Z}_{t+k}$. As $\overline{x}_Z$ is differentiable in each of these terms we obtain

$$\begin{aligned}\left(\int \overline{x} d\Omega\right)_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right) &= \int \mathring{\overline{x}}_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right) d\Omega^* + \int \overline{x}_Z \cdot \hat{Z}_t d\hat{\Omega}_{t+k} + \int \overline{x}_Z \cdot \hat{Z}_{t+k} d\hat{\Omega}_t \\ &\quad - \int \left(\overline{x}_Z^{\Delta}(\theta) \cdot \hat{Z}_t\right) \overline{\kappa}_Z(\theta) \cdot \hat{Z}_{t+k} \omega^*(\overline{\kappa}(\theta), \theta) d\theta \\ &= \int \overline{x}_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right) d\Omega^* + \int \overline{x}_Z \cdot \hat{Z}_t d\hat{\Omega}_{t+k} + \int \overline{x}_Z \cdot \hat{Z}_{t+k} d\hat{\Omega}_t\end{aligned}$$

where the second line is achieved by noting $\overline{x}_Z^{\Delta}(\theta) \cdot \hat{Z}_t = -\overline{x}_a^{\Delta}(\theta)\overline{\kappa}_Z(\theta) \cdot \hat{Z}_{t+k}$ using the formula for the generalized derivative $\overline{x}_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right)$ is defined as the generalized derivative in Claim 5. Adding to it $\left(\int \overline{x} d\Omega\right)_Z \cdot \hat{Z}_{t,t+k} = \int \overline{x}_Z \cdot \hat{Z}_{t,t+k} d\Omega^* + \int \overline{x} d\hat{\Omega}_{t,t+k}$ yields (40).

For brevity, in the remainder of the proofs we will omit the limiting arguments with $\overline{\kappa}(\theta, Z)$ and instead directly use the generalized derivatives.

B.9 Proof of Lemma 5

(a) To determine $\hat{x}_{t,t+k}(a,\theta)$ at points away from the kinks we start with the derivative of the F mapping in the direction $\hat{Z}_{t,t+k}$ and then add to it the second derivative of the F mapping in directions $\hat{Z}_t$ and $\hat{Z}_{t+k}$. Doing so yields

$$0 = \mathsf{F}_x(a,\theta)\hat{x}_{t,k}(a,\theta) + \mathsf{F}_Y(a,\theta)\hat{Y}_{t,t+k} + \mathsf{F}_{x^e}(a,\theta)\left(\left(\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right]\right)_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right) + \left(\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right]\right)_Z \cdot \hat{Z}_{t,t+k}\right)$$
$$+ \ldots.$$

where ... represent interactions of first order terms and are given in $\mathsf{F}_{t,t+k}$ below. The terms $\left(\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right]\right)_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right)$ and $\left(\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right]\right)_Z \cdot \hat{Z}_{t,t+k}$ are obtained by differentiating $\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right] = \int \overline{x}\left(\overline{a}(a,\theta,Z), \theta', \overline{Z}(Z)\right)\mu(\theta' - \rho_\theta\theta)d\theta'$ twice in directions $\hat{Z}_t$ and $\hat{Z}_k$ and once in the direction $\hat{Z}_{t,t+k}$, respectively. This implies their sum is

$$\mathbb{E}_\varepsilon\left[\overline{x}_{aa}|a,\theta\right]\hat{a}_t(a,\theta)\hat{a}_{t+k}(a,\theta) + \mathbb{E}\left[\hat{x}_{aZ,t+k+1}|a,\theta\right]\hat{a}_t(a,\theta) + \mathbb{E}\left[\hat{x}_{aZ,t+1}|a,\theta\right]\hat{a}_{t+k}(a,\theta) + \mathbb{E}_\varepsilon\left[\hat{x}_{t,t+k}|a,\theta\right].$$

Therefore, for points (a,θ) not on the kinks the classical derivative, $\mathring{\hat{x}}_{t,t+k} := \mathring{\overline{x}}_{ZZ} \cdot (\hat{Z}_t, \hat{Z}_{t+k}) + \overline{x}_Z \cdot \hat{Z}_{t,t+k}$ satisfies

$$\mathsf{F}_x(a,\theta)\mathring{\hat{x}}_{t,t+k}(a,\theta) + \mathsf{F}_Y(a,\theta)\hat{Y}_{t,t+k} + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\mathring{\hat{x}}_{t+1,t+k+1}|a,\theta\right] + \mathsf{F}_{t,k}(a,\theta) = 0 \tag{76}$$

with

$$
\begin{aligned}
\mathsf{F}_{t,t+k}(a,\theta) = \mathsf{F}_{x^e}(a,\theta) \Bigg(& \mathbb{E}_\varepsilon\left[\overline{x}_{aa}|a,\theta\right]\hat{a}_t(a,\theta)\hat{a}_{t+k}(a,\theta) + \mathbb{E}_\varepsilon\left[\hat{x}_{aZ,t+k+1}|a,\theta\right]\hat{a}_t(a,\theta) \\
& + \mathbb{E}_\varepsilon\left[\hat{x}_{aZ,t+1}|a,\theta\right]\hat{a}_{t+k}(a,\theta) + \mathbb{E}_\varepsilon\left[\hat{x}^\delta_{t,t+k}|a,\theta\right] \Bigg) \\
& + \mathsf{F}_{xx}(a,\theta)\cdot\left(\hat{x}_t(a,\theta),\hat{x}_{t+k}(a,\theta)\right) + \mathsf{F}_{xY}(a,\theta)\cdot\left(\hat{x}_t(a,\theta),\hat{Y}_{t+k}\right) + \mathsf{F}_{xx^e}(a,\theta)\cdot\left(\hat{x}_t(a,\theta),\hat{x}^e_{t+k}(a,\theta)\right) \\
& + \mathsf{F}_{Yx}(a,\theta)\cdot\left(\hat{Y}_t,\hat{x}_{t+k}(a,\theta)\right) + \mathsf{F}_{YY}(a,\theta)\cdot\left(\hat{Y}_t,\hat{Y}_{t+k}\right) + \mathsf{F}_{Yx^e}(a,\theta)\cdot\left(\hat{Y}_t,\hat{x}^e_{t+k}(a,\theta)\right) \\
& + \mathsf{F}_{x^e x}(a,\theta)\cdot\left(\hat{x}^e_t(a,\theta),\overline{x}_{Z,k}(a,\theta)\right) + \mathsf{F}_{x^e Y}(a,\theta)\cdot\left(\hat{x}^e_t(a,\theta),\hat{Y}_{t+k}\right) + \mathsf{F}_{x^e x^e}(a,\theta)\cdot\left(\hat{x}^e_t(a,\theta),\hat{x}^e_{t+k}(a,\theta)\right)
\end{aligned}
$$

where $\hat{x}^\delta_{t,t+k}(a,\theta) := \hat{x}_{t,t+k} - \mathring{\hat{x}}_{t,t+k} = \overline{x}^\Delta_a(\theta)\hat{\kappa}_t(\theta)\hat{\kappa}_{t+k}(\theta)\delta(a-\overline{\kappa}(\theta))$ and $\hat{x}^e_t(a,\theta) := \mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{p}\hat{x}_t(a,\theta) + \mathbb{E}_\varepsilon\left[\hat{x}_{t+1}|a,\theta\right]$. All of these objects easily constructed from first order terms.

Finally, if we subtract off (43) from (76) we find

$$
\mathsf{F}_x(a,\theta)\mathring{\hat{x}}_{t,t+k}(a,\theta) - \mathring{\mathsf{x}}_{t,t+k}(a,\theta) + \mathsf{F}_Y(a,\theta)\overline{Y}_{ZZ,t,k} + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\mathring{\hat{x}}_{t+1,t+1+k} - \mathring{\mathsf{x}}_{t+1,t+1+k}|a,\theta\right] = 0
$$

which, following the same steps as the proof of Lemma 3, implies

$$
\mathring{\hat{x}}_{t,t+k} - \mathring{\mathsf{x}}_{t,t+k} = \sum_{s=0}^{\infty}\mathsf{x}_s\hat{Y}_{t+s,t+k+s}
$$

and completes the proof.

(b) Assume knowledge of $\overline{x}_{ZZ,0,0}(a,\theta)$. To find $\hat{x}_{\sigma\sigma,t}(a,\theta)$, for any (a,θ) not on a kink, differentiate the F mapping twice with respect to σ and add to it the derivative of F in direction $\hat{Z}_{\sigma\sigma,t}$

$$
0 = \mathsf{F}_x(a,\theta)\hat{x}_{\sigma\sigma,t}(a,\theta) + \mathsf{F}_Y(a,\theta)\hat{Y}_{\sigma\sigma,t} + \mathsf{F}_{x^e}(a,\theta)\left(\left(\mathbb{E}_{\varepsilon,\mathcal{E}}\left[\overline{x}|a,\theta\right]\right)_{\sigma\sigma} + \left(\mathbb{E}_{\varepsilon,\mathcal{E}}\left[\overline{x}|a,\theta\right]\right)_Z\cdot\hat{Z}_{\sigma\sigma,t}\right)
$$

where $\mathbb{E}_{\varepsilon,\mathcal{E}}\left[\overline{x}|a,\theta\right] = \iint \overline{x}\left(\overline{a}(a,\theta,Z;\sigma),\theta',\overline{Z}(\sigma\mathcal{E},Z;\sigma)\right)\mu\left(\theta'-\rho_\theta\theta\right)d\theta' d\Pr\left(\mathcal{E}\right)$ and $\overline{Z}(\sigma\mathcal{E},Z;\sigma) = \left[\sigma\mathcal{E},\mathsf{P}\overline{X}(Z;\sigma),\overline{\Omega}(Z;\sigma)\right]^T$. Taking the second derivative of this object with respect to σ and adding to it the derivative in direction $\hat{Z}_{\sigma\sigma,t}$ yields[34]

$$
\left(\mathbb{E}_{\varepsilon,\mathcal{E}}\tilde{x}\right)_{\sigma\sigma,t} = \mathbb{E}_\varepsilon\left[\hat{x}_{0,0}|a,\theta\right]\mathrm{var}(\mathcal{E}) + \mathbb{E}_\varepsilon\left[\hat{x}_{\sigma\sigma,t+1}|a,\theta\right] + \mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{P}\hat{x}_{\sigma\sigma,t}(a,\theta)
$$

with $\hat{x}_{0,0}$ representing the distributional derivative in Section B.2. Let $\mathsf{x}_{\sigma\sigma}(a,\theta)$ be the function that solves the following linear functional equation

$$
0 = \mathsf{F}_x(a,\theta)\mathsf{x}_{\sigma\sigma}(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\overline{x}_{ZZ,0,0}|a,\theta\right]\mathrm{var}(\mathcal{E}) + \mathbb{E}_\varepsilon\left[\mathsf{x}_{\sigma\sigma}|a,\theta\right] + \mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{x}_{\sigma\sigma}(a,\theta)\right).
$$

Subtracting these two equations we see that $d\hat{x}_{\sigma\sigma,t} := \hat{x}_{\sigma\sigma,t} - \mathsf{x}_{\sigma\sigma}$ satisfies

$$
0 = \mathsf{F}_x(a,\theta)d\hat{x}_{\sigma\sigma,t}(a,\theta) + \mathsf{F}_Y(a,\theta)\overline{X}_{\sigma\sigma,t} + \mathsf{F}_{x'}(a,\theta)\left(\mathbb{E}\left[d\hat{x}_{\sigma\sigma,t+1}|a,\theta\right] + \mathbb{E}\left[\overline{x}_a|a,\theta\right]\mathsf{p}d\hat{x}_{\sigma\sigma,t}(a,\theta)\right).
$$

This is identical to system of equations solved by $\hat{x}_t$ which allows us to conclude that $d\hat{x}_{\sigma\sigma,t} = \sum_{s=0}^{\infty}\mathsf{x}_s(a,\theta)\overline{X}_{\sigma\sigma,t+s}$ which implies (44).

[34]While, $\hat{x}_{ZZ,0,0}$ is a generalized function, $\hat{x}_{\sigma\sigma,t}$ can just be replaced with a classical derivative. Since, $\overline{x}_\sigma$ is uniformly zero, the same steps that show distributional derivatives are equal to the classical derivatives to first order can be used to show $\overline{x}_{\sigma\sigma} = \mathring{\overline{x}}_{\sigma\sigma}$.

B.10 Proof of Lemma 6

(a) Start by differentiating the LoM in direction $\hat{Z}_{t,t+k}$. The same arguments as the first order gives

$$\frac{d}{d\theta}\overline{\Omega}_Z \cdot \hat{Z}_{t,t+k} = \mathcal{L}^{(a)} \cdot \frac{d}{d\theta}\hat{\Omega}_{t,t+k} - \mathcal{M} \cdot \left(\overline{a}_Z \cdot \hat{Z}_{t,t+k}\right).$$

To get $\frac{d}{d\theta}\hat{\Omega}_{t+1,t+1+k}$ we add to it the derivative in the direction $\hat{Z}_{t+k}$ of the expression below

$$\begin{aligned}\frac{d}{d\theta}\overline{\Omega}_Z(Z) \cdot \hat{Z}_t\langle a',\theta'\rangle = &\int\int \overline{\Lambda}(a',\theta',a,\theta,Z)\overline{a}_a(a,\theta,Z)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle dad\theta \\ &- \int\int \overline{\Lambda}(a',\theta',a,\theta,Z)\overline{a}_Z(a,\theta,Z)\cdot\hat{Z}_t d\Omega\langle a,\theta\rangle\end{aligned}$$

where $\overline{\Lambda}(a',\theta',a,\theta,Z) \equiv \delta(\overline{a}(a,\theta,Z)-a')\mu(\theta'-\rho_\theta\theta)$. As

$$\begin{aligned}\overline{\Lambda}_Z(a',\theta',a,\theta)\cdot\hat{Z}_{t+k} &= \delta'(\overline{a}(a,\theta)-a')\mu(\theta'-\rho_\theta\theta)\overline{a}_Z(a,\theta)\cdot\hat{Z}_{t+k} \\ &= -\frac{d}{da'}\delta(\overline{a}(a,\theta)-a')\mu(\theta'-\rho_\theta\theta)\hat{a}_{t+k}(a,\theta) = -\frac{d}{da'}\overline{\Lambda}(a',\theta',a,\theta)\hat{a}_{t+k}(a,\theta)\end{aligned}$$

we have

$$\begin{aligned}\frac{d}{d\theta}\hat{\Omega}_{t+1,t+k+1}\langle a',\theta'\rangle = &\iint \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\frac{d}{d\theta}\hat{\Omega}_{t,t+k}\langle a,\theta\rangle dad\theta - \iint \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_{t,t+k}(a,\theta)d\Omega^* \\ &+ \iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\hat{a}_{aZ,t+k}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle dad\theta - \frac{d}{da'}\iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\overline{a}_a(a,\theta)\hat{a}_{t+k}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle dad\theta \\ &- \iint \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_t(a,\theta)d\hat{\Omega}_{t+k} + \frac{d}{da'}\iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\hat{a}_{t+k}(a,\theta)\hat{a}_t(a,\theta)d\Omega^*\end{aligned}$$

Finally, integration by parts implies

$$\begin{aligned}\iint \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{Z,t}(a,\theta)d\hat{\Omega}_{t+k} = &-\iint \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_{aZ,t+k}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle dad\theta \\ &+ \frac{d}{da'}\iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\overline{a}_a(a,\theta)\hat{a}_t(a,\theta)\frac{d}{d\theta}\hat{\Omega}_{t+k}\langle a,\theta\rangle dad\theta\end{aligned}$$

All combined, using distributional derivatives to absorb the derivatives of the kinks, we have

$$\begin{aligned}\frac{d}{d\theta}\hat{\Omega}_{t+1,t+k+1}\langle a',\theta'\rangle = &\iint \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\frac{d}{d\theta}\hat{\Omega}_{t,t+k}\langle a,\theta\rangle dad\theta - \iint \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_{t,t+k}(a,\theta)d\Omega^* \\ &+ \iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\hat{a}_{aZ,t+k}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle dad\theta - \frac{d}{da'}\iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\overline{a}_a(a,\theta)\hat{a}_{Z,t+k}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle dad\theta \\ &- \iint \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_{aZ,t}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_{t+k}\langle a,\theta\rangle dad\theta + \frac{d}{da'}\iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\overline{a}_a(a,\theta)\hat{a}_t(a,\theta)\frac{d}{d\theta}\hat{\Omega}_{t+k}\langle a,\theta\rangle dad\theta \\ &+ \frac{d}{da'}\iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\hat{a}_{t+k}(a,\theta)\hat{a}_t(a,\theta)d\Omega^*\end{aligned}$$

which can be written more concisely as (46).

(b) Start by differentiating the LoM in direction $\hat{Z}_{\sigma\sigma,t}$. The same arguments as the first order gives $\frac{d}{d\theta}\overline{\Omega}_Z\cdot\hat{Z}_{\sigma\sigma,t} = \mathcal{L}^{(a)}\cdot\hat{\Omega}_{\sigma\sigma,t} - \mathcal{M}\cdot\left(\overline{a}_Z\cdot\hat{Z}_{\sigma\sigma,t}\right)$. Next we take second derivative of (27) with respect to σ. As all the first derivatives w.r.t σ are zero we can follow the same steps as the first order to obtain $\frac{d}{d\theta}\overline{\Omega}_{\sigma\sigma} = -\mathcal{M}\cdot\overline{a}_{\sigma\sigma}$. As $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t+1} = \frac{d}{d\theta}\overline{\Omega}_Z\cdot\hat{Z}_{\sigma\sigma,t} + \frac{d}{d\theta}\overline{\Omega}_{\sigma\sigma}$, adding these two equations together completes the proof.

B.11 Proof of Corollary 2

(a) We begin with the following claim regarding $\frac{d}{da}$ and $\mathcal{L}^{(a)}$

Claim 9. Suppose that y is a generalized function with a finite number of mass points at $\{a_n^*\}$ then: $\mathcal{L}^{(a)} \cdot \frac{d}{da}\mathsf{y} = -\mathcal{L}^{(aa)} \cdot \mathsf{y} + \frac{d}{da}\mathcal{L}^{(a,a)} \cdot \mathsf{y}$.

Proof. We have

$$\left(\mathcal{L}^{(a)} \cdot \frac{d}{da}\mathsf{y}\right)(a^{'},\theta^{'}) = \int\int \overline{\Lambda}(a^{'},\theta^{'},a,\theta)\overline{a}_a(a,\theta)\frac{d}{da}\mathsf{y}(a,\theta)dad\theta$$
$$= -\int \overline{\Lambda}(a^{'},\theta^{'},a,\theta)\overline{a}_{aa}(a,\theta)\mathsf{y}(a,\theta)dad\theta - \int \overline{\Lambda}_a(a^{'},\theta^{'},a,\theta)\overline{a}_a(a,\theta)\mathsf{y}(a,\theta)dad\theta$$
$$= -\int \underbrace{\overline{\Lambda}(a^{'},\theta^{'},a,\theta)\overline{a}_{aa}(a,\theta)\mathsf{y}(a,\theta)dad\theta}_{:=\left(\mathcal{L}^{(aa)}\cdot\mathsf{y}\right)(a',\theta')} + \frac{d}{da'}\underbrace{\int \overline{\Lambda}(a^{'},\theta^{'},a,\theta)\overline{a}_a(a,\theta)\overline{a}_a(a,\theta)\mathsf{y}(a,\theta)dad\theta}_{:=\left(\mathcal{L}^{(a,a)}\cdot\mathsf{y}\right)(a',\theta')}$$

where the second equality uses integration by parts and the third equality uses $\overline{\Lambda}_a(a^{'},\theta^{'},a,\theta) = -\frac{d}{da'}\overline{\Lambda}(a^{'},\theta^{'},a,\theta)\overline{a}_a(a,\theta)$. That y is a generalized function with a finite number of mass points at $\{a_n^*\}$ guarantees that these integrals are well defined using the same arguments as the first order. □

Claim 9 allows us to prove the following claim on $\frac{d}{d\theta}\hat{\Omega}_{t,t+k}$

Claim 10. $\frac{d}{d\theta}\hat{\Omega}_{t,k}$ is given by $\frac{d}{d\theta}\hat{\Omega}_{t,t+k} = -\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\hat{Y}_{s,k+s} - \mathsf{B}_{t,t+k} + \frac{d}{da}\mathsf{C}_{t,t+k}$ where $\mathsf{A}_{t,s}$ is as defined in Claim 8, and $\mathsf{B}_{t,k}$ and $\mathsf{C}_{t,k}$ are defined in Corollary 2.

Proof. The case when $t = 0$ is trivial as $\hat{\Omega}_{0,k} = 0$ and $\mathsf{A}_{0,s} = \mathsf{B}_{0,k} = \mathsf{C}_{0,k}$. We then proceed by induction using the LoM

$$\frac{d}{d\theta}\hat{\Omega}_{t+1,k+1} = \mathcal{L}^{(a)} \cdot \left(-\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\hat{Y}_{s,k+s} - \mathsf{B}_{t,t+k} + \frac{d}{da}\mathsf{C}_{t,t+k}\right) - \sum_{s=0}^{\infty}\mathsf{a}_{s-t}\hat{Y}_{s,t+s} - \mathsf{b}_{t,t+k} + \frac{d}{da}\mathsf{c}_{t,t+k}$$
$$= -\sum_{s=0}^{\infty}\left(\mathcal{L}^{(a)} \cdot \mathsf{A}_{t,s} + \mathsf{a}_{s-t}\right)\hat{Y}_{s,k+s} - \left(\mathcal{L}^{(a)} \cdot \mathsf{B}_{t,t+k} + \mathsf{b}_{t,t+k}\right) - \mathcal{L}^{(aa)} \cdot \mathsf{C}_{t,t+k} + \frac{d}{da}\mathcal{L}^{(a,a)} \cdot \mathsf{C}_{t,t+k} + \frac{d}{da}\mathsf{c}_{t,k}$$
$$= -\sum_{s=0}\mathsf{A}_{t+1,s}\hat{Y}_{s,k+s} - \mathsf{B}_{t+1,t+k+1} + \frac{d}{da}\mathsf{C}_{t+1,t+k+1}$$

as desired. □

Next, by applying integration by parts we have

$$\int \overline{x}d\hat{\Omega}_{t,t+k} = -\iint \overline{x}_a\left(-\sum_{s=0}\mathsf{A}_{t,s}\hat{Y}_{s,k+s} - \mathsf{B}_{t,t+k} + \frac{d}{da}\mathsf{C}_{t,t+k}\right)dad\theta$$
$$= \sum_{s=0}^{\infty}\left(\mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s}\right)\hat{Y}_{s,k+s} + \mathcal{I}^{(a)} \cdot \mathsf{B}_{t,t+k} - \iint \overline{x}_a(a,\theta)\frac{d}{da}\mathsf{C}_{t,t+k}(a,\theta)dad\theta$$
$$= \sum_{s=0}^{\infty}\left(\mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s}\right)\hat{Y}_{s,k+s} + \mathcal{I} \cdot \mathsf{B}_{t,t+k} + \underbrace{\int \overline{x}_{aa}\mathsf{C}_{t,t+k}dad\theta}_{:=\mathcal{I}^{(aa)}\cdot\mathsf{C}_{t,t+k}}$$

We can use the same arguments as the first order to guarantee that both $\mathsf{B}_{t,t+k}$ and $\mathsf{C}_{t,t+k}$ are generalized functions with a finite number of mass points at $\{a_n^*\}$, which guarantees the integrals are well defined.

Finally, tuning to equation (40), we note that

$$\int \hat{x}_k d\hat{\Omega}_t = \iint \hat{x}_k(a,\theta) d\hat{\Omega}_t\langle a,\theta\rangle = -\underbrace{\iint \hat{x}_{aZ,k}\frac{d}{d\theta}\hat{\Omega}_t da d\theta}_{:=\mathcal{I}^{(a)}_{Z,k}\cdot\frac{d}{d\theta}\hat{\Omega}_t}$$

and similarly for $\int \overline{x}_{Z,t} d\hat{\Omega}_k$. Substituting all of these results into equation (40) yields the result of Corollary 2.

(b) We begin with the following Claim

Claim 11. $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}$ is given by $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t} = -\sum_{s=0}\mathsf{A}_{t,s}\hat{Y}_{\sigma\sigma,s} - \mathsf{B}_{\sigma\sigma,t}$, where $\mathsf{A}_{t,s}$ is as defined in Corollary 1 and $\mathsf{B}_{\sigma\sigma,0} = 0$ and $\mathsf{B}_{\sigma\sigma,t+1} = \mathcal{L}^{(a)}\cdot\mathsf{B}_{\sigma\sigma,t} + \mathcal{M}\cdot\mathsf{px}_{\sigma\sigma}$.

Proof. It's trivially true for $t = 0$ as $\mathsf{A}_{0,s} = 0, \mathsf{B}_{\sigma\sigma,0} = 0$, and $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,0} = 0$. We then proceed by induction using the LoM

$$\begin{aligned}\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t+1} &= \mathcal{L}^{(a)}\cdot\left(-\sum_{s=0}\mathsf{A}_{t,s}\hat{Y}_{\sigma\sigma,s} - \mathsf{B}_{\sigma\sigma,t}\right) - \mathcal{M}\cdot\left(\sum_{s=0}^{\infty}\mathsf{px}_{t-s}\hat{Y}_{\sigma\sigma,s} + \mathsf{px}_{\sigma\sigma}\right)\\ &= -\sum_{s=0}^{\infty}\left(\mathcal{L}^{(a)}\cdot\mathsf{A}_{t,s} + \mathcal{M}\cdot\mathsf{px}_{t-s}\right) - \left(\mathcal{L}^{(a)}\cdot\mathsf{B}_{\sigma\sigma,t} + \mathcal{M}\cdot\mathsf{px}_{\sigma\sigma}\right)\end{aligned}$$

which completes the proof. □

Integration by parts then implies that

$$\int \overline{x} d\hat{\Omega}_{\sigma\sigma,t} = -\int \overline{x}_a \frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t} da d\theta = \sum_{s=0}^{\infty}\left(\mathcal{I}^{(a)}\cdot\mathsf{A}_{t,s}\right)\hat{Y}_{\sigma\sigma,s} + \mathcal{I}^{(a)}\cdot\mathsf{B}_{\sigma\sigma,t}$$

where the same arguments as in the first-order guarantee that $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}$ is a generalized function with mass-points only at $\{a^*_n\}$ and thus the operation $\mathcal{I}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}$ is well defined. This implies that

$$\begin{aligned}\left(\widehat{\int \overline{x} d\Omega}\right)_{\sigma\sigma,t} &= \int \overline{x}_{\sigma\sigma,t} d\Omega^* + \int \overline{x} d\hat{\Omega}_{\sigma\sigma,t}\\ &= \sum_{s=0}^{\infty}\underbrace{\left(\int \mathsf{x}_{s-t} d\Omega^* + \mathcal{I}^{(a)}\cdot\mathsf{A}_{t,s}\right)}_{\mathsf{J}_{t,s}}\hat{Y}_{\sigma\sigma,s} + \underbrace{\int \mathsf{x}_{\sigma\sigma} d\Omega^* + \mathcal{I}^{(a)}\cdot\mathsf{B}_{\sigma\sigma,t}}_{:=\mathsf{H}_{\sigma\sigma,t}}\end{aligned}$$

the LoM for $\mathsf{B}_{\sigma\sigma,t}$ in Claim 11 implies that $\mathsf{H}_{\sigma\sigma,t}$ satisfies the recursion $\mathsf{H}_{\sigma\sigma,0} = \int \mathsf{x}_{\sigma\sigma} d\Omega^*$ and $\mathsf{H}_{\sigma\sigma,t+1} = \mathsf{H}_{\sigma\sigma,t} + \mathcal{I}^{(a)}\cdot\left(\mathcal{L}^{(a)}\right)^{t-1}\cdot\mathcal{M}\cdot\mathsf{px}_{\sigma\sigma}$, which completes the proof.

B.12 Proof of Proposition 2^{HA}

The Proposition is a direct result of Corollary 2.

C Comparison to Approximating the Distribution with A Histogram

All the terms in this section will implicitly index everything by h: the space between points along each dimension a and θ. We let $a_{[i]}$ be the gridpoints along the a dimension and $\theta_{[j]}$ be the grid points along the asset dimension. To construct the histogram approach we define projection function $\mathcal{P}^{i,j}(a,\theta)$ as the probability of assigning point a,θ to gridpoint $(a,\theta)_{[i,j]}$. Following Young (2010) we project to the closest neighbors: $\mathcal{P}^{i,j}(a,\theta) = \mathcal{P}^{i}(a)\mathcal{Q}^{j}(\theta)$, where $\mathcal{P}^{i}(a) = \begin{cases} \frac{a-a_{[i]}}{h} & a \in [a_{[i]}, a_{[i+1]}] \\ \frac{a_{[i]}-a}{h} & a \in [a_{[i-1]}, a_{[i]}] \\ 0 & \text{otherwise} \end{cases}$ and similarly for $\mathcal{Q}^{j}(\theta)$.

We assume full knowledge of $\tilde{x}(a,\theta,Z)$ and focus purely on the approximation with respect to the histogram. The approximation to the steady state transition density is $\overline{\Lambda}(i',j',a,\theta) = \int \mathcal{P}^{i',j'}\left(\overline{a}(a,\theta), \rho_\theta\theta + \varepsilon\right) d\mu(\epsilon)$. This constructs a steady state transition matrix $\overline{\Lambda}(i',j',i,j) = \overline{\Lambda}(i',j',a_{[i]},\theta_{[j]})$. We let $\omega^*_{[i,j]}$ be the approximation to the steady state density. We assume that all of these objects are well approximated as $h \to 0$ so for any smooth test function $\phi(a,\theta)$, $\int \phi d\Omega^* = \lim_{h\to 0}\sum_{i,j}\phi\left(a_{[i]},\theta_{[j]}\right)\omega^*_{[i,j]}$ and

$$\int \phi(a',\theta')\overline{\Lambda}(a',\theta',a,\theta)da'd\theta' = \lim_{h\to 0}\sum_{i',j'}\phi\left(a_{[i']},\theta_{[j']}\right)\overline{\Lambda}(i',j',a,\theta).$$

Given $\overline{a}(a,\theta,Z)$, the approximated LoM for the distribution is

$$\overline{a}_{[i',j']}(Z) = \sum_{i,j}\int \mathcal{P}^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]},Z), \rho_\theta\theta + \varepsilon\right) d\mu(\epsilon)\omega_{[i,j]}$$

Differentiating with respect to Z in direction $\hat{Z}$ yields

$$\hat{\omega}_{t+1,[i',j']} = \sum_{i,j}\overline{\Lambda}\left(i',j',i,j\right)\hat{\omega}_{t,[i,j]} + \sum_{i,j}\int \mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}), \rho_\theta\theta + \varepsilon\right) d\mu(\epsilon)\hat{a}_t\left(a_{[i]},\theta_{[j]}\right)\omega^*_{[i,j]}$$

Which we can write succinctly as $\hat{\omega}_{t+1} = \overline{\Lambda}\hat{\omega}_t + \mathcal{M}^h\overrightarrow{a}_t$, where $\overrightarrow{a}_t$ is $\hat{a}_t$ evaluated at the grid-points. With this we prove the following two claims

Claim 12. In the limit as $h \to 0$, $\lim_{h\to 0}\sum_{i',j'}\overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\mathcal{M}^h\overrightarrow{a}_t\right)_{[i',j']} = \mathcal{I}^{(a)}\cdot\mathcal{M}\cdot\hat{a}_t$.

Proof. Note that

$$\sum_{i',j'}\overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\mathcal{M}^h\overrightarrow{a}_t\right)_{[i',j']} = \sum_{i,j}\int\sum_{i',j'}\overline{x}(a_{[i']},\theta_{[j']})\mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}), \rho_\theta\theta + \varepsilon\right) d\mu(\epsilon)\hat{a}_t\left(a_{[i]},\theta_{[j]}\right)\omega^*_{[i,j]}$$

This simplifies as

$$\begin{aligned}\sum_{i',j'}\overline{x}(a_{[i']},\theta_{[j']})\mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}), \rho_\theta\theta + \varepsilon\right) &= \overline{x}_a(\overline{a}(a_{[i]},\theta_{[j]}), \rho_\theta\theta + \varepsilon) + \mathcal{O}(h) \\ &= \sum_{i',j'}\overline{x}_a(a_{[i']},\theta_{[j']})\mathcal{P}^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}), \rho_\theta\theta + \varepsilon\right) + \mathcal{O}(h)\end{aligned}$$

and thus $\sum_{i',j'}\overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\mathcal{M}^h\overrightarrow{a}_t\right)_{[i',j']} = \sum_{i',j'}\sum_{i,j}\overline{x}_a(a_{[i']},\theta_{[j']})\overline{\Lambda}\left(i',j',i,j\right)\hat{a}_t\left(a_{[i]},\theta_{[j]}\right)\omega^*_{[i,j]}$. Taking limit as $h \to 0$ completes the result. ☐

Claim 13. In the limit as $h \to 0$, $\lim_{h\to 0}\sum_{i',j'}\overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\overline{\Lambda}\mathcal{M}^h\overrightarrow{a}_t\right)_{[i',j']} = \mathcal{I}^{(a)}\cdot\mathcal{L}^{(a)}\cdot\mathcal{M}\cdot\hat{a}_t$.

Proof. For this we're going to use that for any smooth function $\phi(a)$

$$\sum_{i'',i'} \phi(a_{[i'']})\mathcal{P}^{i''}(\overline{a}(a_{[i']},\theta))\mathcal{P}^{i'}_a(a) = \frac{1}{h}\sum_{i''}\phi(a_{[i'']})\left(\mathcal{P}^{i''}(\overline{a}(a_{[\hat{i}]}+h,\theta)) - \mathcal{P}^{i''}(\overline{a}(a_{[\hat{i}]},\theta))\right)$$
$$= \phi_a(\overline{a}(a,\theta))\overline{a}_a(a,\theta) + \mathcal{O}(h)$$

We therefore have that

$$\sum_{i',j'} \overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\overline{\Lambda}\mathcal{M}^h\overrightarrow{a}_t\right)_{[i',j']} = \sum_{i'',j'',i',j'} \overline{x}(a_{[i'']},\theta_{[j'']})\overline{\Lambda}(i'',j'',i',j')\sum_{i,j}\int \mathcal{P}^{i',j'}_a\left(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon\right)d\mu(\epsilon)\hat{a}_t\left(a_{[i]},\theta_{[j]}\right)\overline{\omega}_{[i,j]}$$

We can then exploit the fact that $\overline{\Lambda}(i'',j'',i',j') = \int \mathcal{Q}^{j''}\left(\rho_\theta\theta_{[j']}+\varepsilon\right)d\mu(\epsilon)\mathcal{P}^{i''}\left(\overline{a}(a_{[i']},\theta_{[j']})\right)$ to get

$$\sum_{i',j'} \overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\overline{\Lambda}\mathcal{M}^h\overrightarrow{a}_t\right)_{[i',j']} = \sum_{i'',j'',i',j'} \overline{x}_a(a_{[i'']},\theta_{[j'']})\overline{\Lambda}(i'',j'',i',j')\overline{a}_a(a_{[i']},\theta_{[j']})\overline{\Lambda}(i'',j'',i',j')\hat{a}_t\left(a_{[i]},\theta_{[j]}\right)\overline{\omega}_{[i,j]}+\mathcal{O}(h)$$

which in the limit as $h\to 0$ gives $\mathcal{I}^{(a)}\cdot\mathcal{L}^{(a)}\cdot\mathcal{M}\cdot\hat{a}_t$. □

This same argument extends to show that $\lim_{h\to 0}\sum_{i',j'}\overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\overline{\Lambda}^t\mathcal{M}^h\overrightarrow{a}_t\right)_{[i',j']} = \mathcal{I}^{(a)}\cdot\left(\mathcal{L}^{(a)}\right)^t\cdot\mathcal{M}\cdot\hat{a}_t$ for arbitrary t.

C.1 Second Order

Taking a second derivative of the LoM when $t=k=0$ we have, after exploiting that $\mathcal{P}^{i,j}_{aa}=0$,

$$\hat{\omega}_{ZZ,1,1,[i',j']} = \sum_{i,j}\int \mathcal{P}^{i',j'}_a\left(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon\right)d\mu(\epsilon)\overline{a}_{ZZ,0,0}\left(a_{[i]},\theta_{[j]}\right)\overline{\omega}_{[i,j]},$$

which implies $\lim_{h\to 0}\sum_{i',j'}\overline{x}(a_{[i']},\theta_{[j']})\hat{\omega}_{ZZ,1,1,[i',j']} = \mathcal{I}^{(a)}\cdot\mathcal{M}\cdot\hat{a}_{0,0}$. As $\int \overline{x}d\hat{\Omega}_{1,1} = \mathcal{I}^{(a)}\cdot\mathcal{M}\cdot\hat{a}_{0,0} + \mathcal{I}^{(aa)}\cdot\mathsf{C}_{1,1}$, we conclude the histogram method misses the $\mathsf{C}_{1,1}$ term.

D Proofs of Section 5

In this section we present the proofs for the extensions presented in Section 5. As needed we will use distributional derivatives in place of the classical derivatives,

D.1 Proofs for Section 5.1

D.1.1 Proof of Proposition 1^{TD}

Let $\hat{Z}^{TD}_0 = \left[0,\hat{A}^{TD}_0,\hat{\Omega}^{TD}_0\right]$ and define the directions $\hat{Z}^{TD}_t = \left[0,\hat{A}^{TD}_t,\hat{\Omega}^{TD}_t\right]$ recursively via $\hat{Z}^{TD}_t = \overline{Z}_Z\cdot\hat{Z}^{TD}_{t-1}$. Define $\hat{X}^{TD}_t$ as $\overline{X}_Z\cdot\hat{Z}^{TD}_t$. We start with the following claim

Claim 14. To the first-order approximation, X_t satisfies $\mathbb{E}_0 X_t = \overline{X}+\hat{X}^{TD}_t + O\left(\left\|\mathcal{E},\hat{Z}_0\right\|^2\right)$.

Proof. The path of aggregates, $X_t(\mathcal{E}^t;\Theta_{-1},A_{-1},\Omega_0,\sigma)$, depends on the history of aggregate shocks, $\mathcal{E}^t$, and the initial state $\Theta_{-1},A_{-1},\Omega_0$. It can be constructed from the recursive representation $\tilde{X}(Z;\sigma)$ and $\tilde{\Omega}(Z;\sigma)$ by defining $Z_t(\mathcal{E}^t;\Theta_{-1},A_{-1},\Omega_0,\sigma) = \left[\Theta_t\left(\mathcal{E}^t;\sigma\right),\Omega_t\left(\mathcal{E}^t;\Theta_{-1},A_{-1},\Omega_0,\sigma\right)\right]^T$ recursively as follows: let $Z_0(\mathcal{E}_0;\Theta_{-1},A_{-1},\Omega_0,\sigma) = [\rho_\Theta\Theta_{-1}+\sigma\mathcal{E}_0,A_{-1},\Omega_0]^T$ and for $t\geq 1$

$$Z_t\left(\mathcal{E}^t;\Theta_{-1},A_{-1},\Omega_0,\sigma\right) = \left[\rho_\Theta\Theta_{t-1}\left(\mathcal{E}^{t-1};\sigma\right)+\sigma\mathcal{E}_t,\tilde{\Omega}\left(Z_{t-1}\left(\mathcal{E}^{t-1};\Theta_{-1},A_{-1},\Omega_0,\sigma\right);\sigma\right)\right], \quad (77)$$

The path of aggregates can then be defined as $X_t(\mathcal{E}^t;\Theta_{-1},A_{-1},\Omega_0,\sigma)=\tilde{X}\left(Z_t\left(\mathcal{E}^t;\Theta_{-1},A_{-1},\Omega_0,\sigma\right);\sigma\right).$

Defining $\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)$ and $\overline{X}_{t,\sigma}\left(\mathcal{E}^t\right)$ as the derivatives of $Z_t\left(\mathcal{E}^t;A_{-1},\Omega_0,\sigma\right)$ and $X_t(\mathcal{E}^t;A_{-1},\Omega_0,\sigma)$ w.r.t σ evaluated at $\sigma=0$,$\Theta_{-1}=0$, $A_{-1}=A^*$, $\Omega_0=\Omega^*$. The same steps as in the proof of Lemma (1) show that $\overline{X}_{t,\sigma}\left(\mathcal{E}^t\right)=\sum_{s=0}^t\hat{X}_{t-s}\mathcal{E}_s$. Next, taking the derivative of $Z_t\left(\mathcal{E}^t;\Theta_{-1},A_{-1},\Omega_0,\sigma\right)$ and $X_t(\mathcal{E}^t;\Theta_{-1},A_{-1},\Omega_0,\sigma)$ w.r.t. $\Theta_{-1},A_{-1},\Omega_0$ in the direction $\hat{Z}_0^{TD}$, we have $\overline{Z}_{0,Z}\left(\mathcal{E}^0\right)\cdot\hat{Z}_0^{TD}=\hat{Z}_0^{TD}$ and for $t\geq 0$

$$\overline{Z}_{t+1,Z}\left(\mathcal{E}^{t+1}\right)\cdot\hat{Z}_0^{TD}=\overline{Z}_Z\cdot\overline{Z}_{t,\Omega}\left(\mathcal{E}^t\right)\cdot\hat{Z}_0^{TD} \tag{78}$$

$$\overline{X}_{t,Z}\left(\mathcal{E}^t\right)\cdot\hat{Z}_0^{TD}=\overline{X}_Z\cdot\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)\cdot\hat{Z}_0^{TD} \tag{79}$$

which implies that $\overline{Z}_{t,Z}\left(\mathcal{E}^t\right)\cdot\hat{Z}_0^{TD}=\hat{Z}_t^{TD}$ and thus$\overline{X}_{t,Z}\left(\mathcal{E}^t\right)\cdot\hat{Z}_0^{TD}=\hat{X}_t^{TD}$. All put together we have $X_t(\mathcal{E}^t;0,A_{-1},\Omega_0)=\sum_{s=0}^t\hat{X}_{t-s}\mathcal{E}_s+\hat{X}_t^{TD}+O\left(\|\mathcal{E},\hat{A}_0^{TD},\hat{\Omega}_0^{TD}\|^2\right)$. Taking expectations completes the proof. □

Following the same steps as the poofs of Lemma 3, and Lemma 4^{HA} it is possible to show that $\frac{d}{d\theta}\hat{\Omega}_{t+1}^{TD}=\mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_t^{TD}-\mathcal{M}\cdot\hat{a}_t^{TD}$, where $\hat{a}_t^{TD}=\mathsf{p}\hat{x}_t^{TD}$ and $\hat{x}_t^{TD}(a,\theta):=\overline{x}_Z(a,\theta)\cdot\hat{Z}_t^{TD}=\sum_{s=0}^{\infty}\mathsf{x}_s(a,\theta)\hat{Y}_{t+s}^{TD}$. Rolling forward the LoM allows us to prove the following claim

Claim 15. $\frac{d}{d\theta}\hat{\Omega}_t^{TD}$ is given by $\frac{d}{d\theta}\hat{\Omega}_t^{TD}=-\sum_{s=0}\mathsf{A}_{t,s}\hat{Y}_s^{TD}-\mathsf{A}_t^{TD}$, where $\mathsf{A}_{t,s}$ is as defined in Corollary 1 where A_t^{TD} satisfies $\mathsf{A}_{t+1}^{TD}=\mathcal{L}^{(a)}\cdot\mathsf{A}_t^{TD}$ and $\mathsf{A}_0^{TD}=-\frac{d}{d\theta}\hat{\Omega}_0^{TD}$.

Proof. The statement is true for $t=0$ as $\mathsf{A}_{0,s}=0$ and $\mathsf{A}_0^{TD}=-\frac{d}{d\theta}\hat{\Omega}_0^{TD}$. We then proceed by induction

$$\begin{aligned}\frac{d}{d\theta}\hat{\Omega}_{t+1}^{TD}&=\mathcal{L}^{(a)}\cdot\left(-\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\hat{Y}_s^{TD}-\mathsf{A}_t^{TD}\right)-\sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{p}\mathsf{x}_{s-t}\hat{Y}_s^{TD}\\&=-\sum_{s=0}^{\infty}\left(\mathcal{L}^{(a)}\cdot\mathsf{A}_{t,s}+\mathcal{M}\cdot\mathsf{a}_{s-t}\right)\hat{Y}_s^{TD}-\mathcal{L}^{(a)}\cdot\mathsf{A}_t^{TD}\equiv-\sum_{s=0}^{\infty}\mathsf{A}_{t+1,s}\hat{Y}_s^{TD}-\mathsf{A}_{t+1}^{TD},\end{aligned}$$

where we follow the convention $\mathsf{x}_k=0$ for $k<0$. □

Finally,

$$\begin{aligned}\left(\int\overline{x}d\Omega\right)_Z\cdot\hat{Z}_t^{TD}&=-\int\overline{x}_a\frac{d}{d\theta}\hat{\Omega}_t^{TD}dad\theta+\sum_{s=0}^{\infty}\int\mathsf{x}_s(a,\theta)d\Omega^*\hat{Y}_{t+s}^{TD}\\&=\sum_s\left(\int\mathsf{x}_{s-t}(a,\theta)d\Omega^*+\mathcal{I}^{(a)}\cdot\mathsf{A}_{t,s}\right)\hat{Y}_s^{TD}+\mathcal{I}^{(a)}\cdot\mathsf{A}_t^{TD}=\sum_s\mathsf{J}_{t,s}\hat{Y}_s^{TD}+\mathsf{J}_t^{TD}.\end{aligned}$$

D.2 Proofs for Section 5.2

We will prove Lemma 2^{SV} and Proposition 2^{SV} in tandem with each other as they need to be shown jointly. Our use of stochastic volatility contains one major departure from section 3: the approximations are around $(\Upsilon,0,A^*,\Omega^*)$ where Υ is stochastic. As Υ only affects the level of risk, only the second derivatives w.r.t. σ will depend on Υ and thus we will represent them as $\overline{\Omega}_{\sigma\sigma}\left(\Upsilon\right),\overline{Z}_{\sigma\sigma}(\Upsilon),\overline{X}_{\sigma\sigma}(\Upsilon)$ and $\overline{x}_{\sigma\sigma}(a,\theta,\Upsilon)$.

Define the stochastic set of directions $\hat{Z}_{\sigma\sigma,t}^{\Upsilon}(\mathcal{E}_{\Upsilon}^t)$ recursively as follows: $\hat{Z}_{\sigma\sigma,0}^{\Upsilon}\left(\mathcal{E}_{\Upsilon}^0\right)=0$ and[35] $\hat{Z}_{\sigma\sigma,t}^{\Upsilon}\left(\mathcal{E}_{\Upsilon}^t\right)=\overline{Z}_Z\hat{Z}_{\sigma\sigma,t-1}^{\Upsilon}(\mathcal{E}_{\Upsilon}^{t-1})+\overline{Z}_{\sigma\sigma}(\Upsilon_{t-1})$. Similarly, define $\hat{X}_{\sigma\sigma,t}^{\Upsilon}\left(\mathcal{E}_{\Upsilon}^t\right)=\overline{X}_Z\cdot\hat{Z}_{\sigma\sigma,t}^{\Upsilon}\left(\mathcal{E}_{\Upsilon}^t\right)+\overline{X}_{\sigma\sigma}\left(\Upsilon_t\right).$ Our first result generalizes Lemma 2 to allow for stochastic volatility

[35]There are also $\overline{X}_{\sigma Z}$ and $\overline{Z}_{\sigma Z}$ terms but they are 0 following the same logic as $\overline{X}_{\sigma}$ and $\overline{Z}_{\sigma}$ being 0 in the proof of Lemma 1

Lemma 2^{SValt}. *To the second-order approximation, X_t satisfies*

$$X_t\left(\boldsymbol{\mathcal{E}}^t\right) = \overline{X} + \sum_{s=0}^{t}\hat{X}_{t-s}\mathcal{E}_s + \frac{1}{2}\left(\sum_{s=0}^{t}\sum_{m=0}^{t}\hat{X}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m\right) \\ + \frac{1}{2}\hat{X}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}^t_{\Upsilon}\right) + O\left(\|\mathcal{E}\|^3\right), \tag{80}$$

where sequences $\{\hat{X}_t\}_t$, $\{\hat{X}_{t,k}\}_{t,k}$ are the same ones as in Propositions 1^{HA} and 2^{HA}.

Proof. We proceed in the same manner as the proof of Lemma 2. Second-order derivatives of (9) and (8) w.r.t. σ to find $\overline{Z}_{0,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^0\right) = 0$ and

$$\overline{Z}_{t+1,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^{t+1}\right) = \overline{Z}_Z\cdot\overline{Z}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right) + \overline{Z}_{ZZ}\cdot\left(\overline{Z}_{t,\sigma}\left(\boldsymbol{\mathcal{E}}^t\right), \overline{Z}_{t,\sigma}\left(\boldsymbol{\mathcal{E}}^t\right)\right) + \overline{Z}_{\sigma\sigma}(\Upsilon_t) \tag{81}$$

$$\overline{X}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right) = \overline{X}_Z\cdot\overline{Z}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right) + \overline{X}_{ZZ}\cdot\left(\overline{Z}_{t,\sigma}\left(\boldsymbol{\mathcal{E}}^t\right), \overline{Z}_{t,\sigma}\left(\boldsymbol{\mathcal{E}}^t\right)\right) + \overline{X}_{\sigma\sigma}(\Upsilon_t) \tag{82}$$

where $\overline{Z}_{ZZ}$ is defined in the main text and $\overline{Z}_{\sigma\sigma} := \left[0, \mathsf{P}\overline{X}_{\sigma\sigma}\left(\Upsilon_t\right), \overline{\Omega}_{\sigma\sigma}(\Upsilon_t)\right]^T$.

Next we show the following claim relating $\overline{Z}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right)$ and our directions $\hat{Z}^{\Upsilon}_{\sigma\sigma,t}(\mathcal{E}^t_{\Upsilon})$

Claim 16. For all t

$$\overline{Z}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right) = \hat{Z}^{\Upsilon}_{\sigma\sigma,t}(\mathcal{E}^t_{\Upsilon}) + \sum_{s=0}^{t}\sum_{m=0}^{t}\hat{Z}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m \tag{83}$$

Proof. We proceed by induction. As $\hat{Z}_{\sigma\sigma,0} = \hat{Z}_{0,0} = 0$ we conclude that equation (8) holds for $t=0$ since $\overline{Z}_{0,\sigma\sigma}\left(\mathcal{E}^0\right) = 0$. Assuming (83) holds for $t-1$ we have

$$\begin{aligned}\overline{Z}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right) &= \overline{Z}_Z\cdot\left(\hat{Z}^{\Upsilon}_{\sigma\sigma,t-1}(\mathcal{E}^{t-1}_{\Upsilon}) + \sum_{s=0}^{t-1}\sum_{m=0}^{t-1}\hat{Z}_{t-1-s,t-1-m}\mathcal{E}_s\mathcal{E}_m\right) + \overline{Z}_{ZZ}\cdot\left(\sum_{s=0}^{t-1}\hat{Z}_{t-1-s}\mathcal{E}_s, \sum_{m=0}^{t-1}\hat{Z}_{t-1-m}\mathcal{E}_m\right) + \overline{Z}_{\sigma\sigma}(\Upsilon_{t-1})\\ &= \hat{Z}^{\Upsilon}_{\sigma\sigma,t}(\mathcal{E}^t_{\Upsilon}) + \sum_{s=0}^{t}\sum_{m=0}^{t}\hat{Z}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m,\end{aligned}$$

where in the second equality we used the fact that $\overline{Z}_{ZZ}$ is a bi-linear mapping and in the third equality we use the recursive definitions of $\hat{Z}^{\Upsilon}_{\sigma\sigma,t}$ and $\hat{Z}_{t,k}$, and $\hat{Z}_{0,0} = 0$. □

Finally we plug in for $\overline{Z}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right)$ and $\overline{Z}_{t,\sigma}\left(\boldsymbol{\mathcal{E}}^t\right)$ in equation (82) to find

$$\begin{aligned}\overline{X}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right) &= \overline{X}_Z\cdot\left(\hat{Z}^{\Upsilon}_{\sigma\sigma,t}(\mathcal{E}^t_{\Upsilon}) + \sum_{s=0}^{t}\sum_{m=0}^{t}\hat{Z}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m\right) + \overline{X}_{ZZ}\cdot\left(\sum_{s=0}^{t}\hat{Z}_{t-s}\mathcal{E}_s, \sum_{m=0}^{t}\hat{Z}_{t-m}\mathcal{E}_m\right)\\ &= \hat{X}^{\Upsilon}_{\sigma\sigma,t}(\mathcal{E}^t_{\Upsilon}) + \sum_{s=0}^{t}\sum_{m=0}^{t}\overline{X}_{ZZ,t-s,t-m}\mathcal{E}_s\mathcal{E}_m.\end{aligned}$$

□

This Lemma uses slightly different directional derivatives than Lemma 2^{SV}. The remainder of this proof will show that $\hat{X}^{\Upsilon}_{\sigma\sigma,t}(\mathcal{E}^t_{\Upsilon})$ can be constructed via the sequence of directional derivatives $\hat{X}^{SV}_{\sigma\sigma,k}$ defined in Proposition 2^{SV}: $\hat{X}^{\Upsilon}_{\sigma\sigma,t}(\mathcal{E}^t_{\Upsilon}) = \hat{X}_{\sigma\sigma,t} + \sum_{s=0}^{t}\hat{X}^{SV}_{\sigma\sigma,t-s}\mathcal{E}_{\Upsilon,s}$, which will complete the proof of Lemma 2^{SV}. We show this linear relationship by first understanding the derivatives $\hat{x}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})$ defined by $\hat{x}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon}) := \overline{x}_Z\left(a,\theta\right)\cdot\hat{Z}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}^t_{\Upsilon}\right) + \overline{x}_{\sigma\sigma}\left(a,\theta,\Upsilon_t\right)$. We can then show the following relationship between $\hat{x}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})$ and $\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,t+s}|\mathcal{E}^t_{\Upsilon}\right]$ where

$$\hat{Y}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}^t_{\Upsilon}\right)=\left[0,\mathsf{P}\hat{X}^{\Upsilon}_{t-1}(\mathcal{E}^{t-1}_{\Upsilon}),\hat{X}^{\Upsilon}_{t}(\mathcal{E}^{t}_{\Upsilon}),\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\hat{X}^{\Upsilon}_{\sigma\sigma,t+1}|\mathcal{E}^t_{\Upsilon}\right]+\hat{X}_{ZZ,0,0}\left(1+\Upsilon_t\right)\mathrm{var}(\mathcal{E}_\Theta)\right]^T.$$

Claim 17. Let $\mathsf{x}^{SV}_{\sigma\sigma}$ be as defined in Proposition 2^{SV} and $\mathsf{x}_{\sigma\sigma}$ be as defined in 5 then for any t

$$\hat{x}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})=\sum_{s=0}^{\infty}\mathsf{x}_s(a,\theta)\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,t+s}|\mathcal{E}^t_{\Upsilon}\right]+\mathsf{x}_{\sigma\sigma}(a,\theta)+\Upsilon_t\mathsf{x}^{SV}_{\sigma\sigma}(a,\theta).$$

Proof. To find $\hat{x}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})$ differentiate the F mapping twice with respect to σ (evaluated at $(\Upsilon,0,A^*,\Omega^*)$) and add to it the derivative of F in direction $\hat{Z}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}^t_{\Upsilon}\right)$ to get

$$\begin{aligned}0&=\mathsf{F}_x(a,\theta)\hat{x}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})+\mathsf{F}_Y(a,\theta)\hat{Y}^{\Upsilon}_{\sigma\sigma,t}(\mathcal{E}^t_{\Upsilon})\\&\quad+\mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\overline{x}_{ZZ,0,0}|a,\theta\right](1+\Upsilon_t)\mathrm{var}(\mathcal{E})+\mathbb{E}_{\varepsilon,\mathcal{E}_\Upsilon}\left[\overline{x}_{\sigma\sigma,t+1}|a,\theta,\mathcal{E}^t_{\Upsilon}\right]+\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{P}\overline{x}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})\right).\end{aligned}$$

Define $d\hat{x}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})=\hat{x}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})-\mathsf{x}_{\sigma\sigma}(a,\theta)-\Upsilon_t\mathsf{x}^{SV}(a,\theta)$, then $d\hat{x}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})$ solves

$$0=\mathsf{F}_x(a,\theta)d\hat{x}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})+\mathsf{F}_Y(a,\theta)\hat{Y}^{\Upsilon}_{\sigma\sigma,t}(\mathcal{E}^t_{\Upsilon})+\mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}\left[d\hat{x}^{\Upsilon}_{\sigma\sigma,t+1}|a,\theta,\mathcal{E}^t_{\Upsilon}\right]+\mathbb{E}\left[\overline{x}_a|a,\theta\right]\mathsf{P}d\hat{x}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})\right).$$

This linear system of equations is identical to the one solved by $\hat{x}_t$ which allows us to conclude that $d\hat{x}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})=\sum_{s=0}^{\infty}\mathsf{x}_s(a,\theta)\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,t+s}|\mathcal{E}^t_{\Upsilon}\right]$, which completes the proof. □

Next we'll characterize the LoM of motion for $\hat{\Omega}_{\sigma\sigma,t}\left(\mathcal{E}^t_{\Upsilon}\right):=\overline{\Omega}_Z\cdot\hat{Z}^{\Upsilon}_{\sigma\sigma,t-1}\left(\mathcal{E}^{t-1}_{\Upsilon}\right)+\overline{\Omega}_{\sigma\sigma}\left(\Upsilon_{t-1}\right)$. Differentiating the LoM twice with respect to σ and adding to it the derivative of the LoM in direction $\hat{Z}_{\sigma\sigma,t}\left(\mathcal{E}^t_{\Upsilon}\right)$, after applying integration by parts, yields $\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_{\sigma\sigma,t+1}\left(\mathcal{E}^{t+1}_{\Upsilon}\right)=\mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}^t_{\Upsilon}\right)-\mathcal{M}\cdot\hat{a}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})$. Substituting for $\hat{a}^{\Upsilon}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_{\Upsilon})$ using Claim 17 immediately obtains the LoM

$$\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_{\sigma\sigma,t+1}\left(\mathcal{E}^{t+1}_{\Upsilon}\right)=\mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}^{t}_{\Upsilon}\right)-\sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{p}x_s\mathbb{E}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,t+s}|\mathcal{E}^t_{\Upsilon}\right]+\mathcal{M}\cdot\mathsf{px}_{\sigma\sigma}+\Upsilon_t\mathcal{M}\cdot\mathsf{px}^{SV}_{\sigma\sigma}.$$

We then are able to prove the following Claim about $\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}^t_{\Upsilon}\right)$

Claim 18. $\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}^t_{\Upsilon}\right)$ satisfies $\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_{\sigma\sigma,t}\right]=-\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,s}\right]-\mathsf{B}_{\sigma\sigma,t}$, and for $k\geq 0$

$$\Delta\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_{\sigma\sigma,\tau+k}|\mathcal{E}^\tau_{\Upsilon}\right]=-\sum_{k=0}^{\infty}\mathsf{A}_{k,s}\Delta\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,\tau+s}|\mathcal{E}^\tau_{\Upsilon}\right]-\mathsf{B}^{SV}_{\sigma\sigma,k}\mathcal{E}_{\Upsilon,\tau}$$

where $\Delta\mathbb{E}_{\mathcal{E}_\Upsilon}\left[Y|\mathcal{E}^\tau_{\Upsilon}\right]:=\mathbb{E}_{\mathcal{E}_\Upsilon}\left[Y|\mathcal{E}^\tau_{\Upsilon}\right]-\mathbb{E}_{\mathcal{E}_\Upsilon}\left[Y|\mathcal{E}^{\tau-1}_{\Upsilon}\right]$, $\mathsf{B}_{\sigma\sigma,t}$ is as defined in Claim 11 and $\mathsf{B}^{SV}_{\sigma\sigma,k}$ is defined by $\mathsf{B}^{SV}_{\sigma\sigma,k+1}=\mathcal{L}^{(a)}\cdot\mathsf{B}^{SV}_{\sigma\sigma,k}+\rho^k_{\Upsilon}\mathcal{M}\cdot\mathsf{px}^{\mathsf{SV}}_{\sigma\sigma}$ with $\mathsf{B}^{SV}_{\sigma\sigma,0}=0$.

Proof. We proceed by induction. The first equation is trivially true for $t=0$ as $\mathsf{A}_{0,s}=0,\mathsf{B}_{\sigma\sigma,0}=0$ and $\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_0=0$. We then proceed by induction (exploiting $\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\Upsilon_t\right]=0$)

$$\begin{aligned}\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_{\sigma\sigma,t+1}\right]&=\mathcal{L}^{(a)}\cdot\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_{\sigma\sigma,t}\right]-\sum_{j=0}^{\infty}\mathcal{M}\cdot\mathsf{px}_j\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,t+j}\right]-\mathcal{M}\cdot\mathsf{px}_{\sigma\sigma}\\&=\sum_{s=0}^{\infty}-\underbrace{\left(\mathcal{L}^{(a)}\cdot\mathsf{A}_{t,s}+\mathcal{M}\cdot\mathsf{px}_{s-t}\right)}_{\mathsf{A}_{t+1,s}}\mathbb{E}_{\mathcal{E}_\Upsilon}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,s}\right]-\underbrace{\left(\mathcal{L}^{(a)}\cdot\mathsf{B}_{\sigma\sigma,t}+\mathcal{M}\cdot\mathsf{px}_{\sigma\sigma}\right)}_{\mathsf{B}_{\sigma\sigma,t+1}}.\end{aligned}$$

For the second equation, note that it holds for $k=0$ as $\Delta\mathbb{E}\left[\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,\tau}|\mathcal{E}_{\Upsilon}^{\tau}\right]=0$ (as $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,\tau}$ is measurable w.r..t $\mathcal{E}_{\Upsilon}^{\tau-1}$) and $\mathsf{A}_{0,s}=0, \mathsf{B}^{SV}_{\sigma\sigma,0}=0$. Taking expectations of then implies

$$\begin{aligned}\Delta\mathbb{E}_{\mathcal{E}_{\Upsilon}}\left[\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_{\sigma\sigma,\tau+k+1}|\mathcal{E}_{\Upsilon}^{\tau}\right] &= \mathcal{L}^{(a)}\cdot\left(-\sum_{s=0}^{\infty}\mathsf{A}_{k,s}\Delta\mathbb{E}_{\mathcal{E}_{\Upsilon}}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,\tau+s}|\mathcal{E}_{\Upsilon}^{\tau}\right]-\mathsf{B}^{SV}_{\sigma\sigma,k}\mathcal{E}_{\Upsilon,\tau}\right)-\sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{px}_{s-k}\Delta\mathbb{E}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,\tau+s}|\mathcal{E}_{\Upsilon}^{\tau}\right]\\ &\quad-\rho_{\Upsilon}^{k}\mathcal{M}\cdot\mathsf{px}^{\mathsf{SV}}_{\sigma\sigma}\mathcal{E}_{\Upsilon,\tau}\\ &=\sum_{s=0}^{\infty}-\underbrace{\left(\mathcal{L}^{(a)}\cdot\mathsf{A}_{k,s}+\mathcal{M}\cdot\mathsf{px}_{s-k}\right)}_{\mathsf{A}_{t+1,s}}\Delta\mathbb{E}_{\mathcal{E}_{\Upsilon}}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,\tau+s}|\mathcal{E}_{\Upsilon}^{\tau}\right]-\underbrace{\left(\mathcal{L}^{(a)}\cdot\mathsf{B}^{SV}_{\sigma\sigma,k}+\rho_{\Upsilon}^{k}\mathcal{M}\cdot\mathsf{px}^{\mathsf{SV}}_{\sigma\sigma}\right)}_{\mathsf{B}^{SV}_{\sigma\sigma,k+1}}\mathcal{E}_{\Upsilon,\tau},\end{aligned}$$

which completes the proof ☐

Differentiating the G mapping twice with respect to σ and adding to it the derivative in direction $\hat{Z}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}_{\Upsilon}^{t}\right)$ yields, after applying integration by parts,

$$\mathsf{G}_{x}\int\hat{x}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}_{\Upsilon}^{t}\right)d\Omega^{*}-\mathsf{G}_{x}\int\overline{x}_{a}\frac{d}{d\theta}\hat{\Omega}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}_{\Upsilon}^{t}\right)dad\theta+\mathsf{G}_{Y}\hat{Y}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}_{\Upsilon}^{t}\right)=0. \tag{84}$$

Taking expectations of both sides and substituting for $\mathbb{E}_{\mathcal{E}_{\Upsilon}}\left[\hat{x}^{\Upsilon}_{\sigma\sigma,t}\right]$ and $\mathbb{E}_{\mathcal{E}_{\Upsilon}}\left[\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}\right]$ using Claims 17 and 18 we have $\mathsf{G}_{x}\sum_{s=0}^{\infty}\mathsf{J}_{t,s}\mathbb{E}_{\mathcal{E}_{\Upsilon}}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,s}\right]+\mathsf{G}_{x}\mathsf{H}_{\sigma\sigma,t}+\mathsf{G}_{Y}\mathbb{E}_{\mathcal{E}_{\Upsilon}}\left[\hat{Y}^{\Upsilon}_{\sigma\sigma,t}\right]=0$, which implies that $\mathbb{E}\left[\hat{X}^{\Upsilon}_{\sigma\sigma,s}\right]$ solves the same system of equations as the $\hat{X}_{\sigma\sigma,t}$ terms in Proposition 2^{HA}.

If we instead take expectations of 19 conditional on $\mathcal{E}_{\Upsilon}^{\tau}$ and subtract off the expectation conditional on $\mathcal{E}_{\Upsilon}^{t-1}$ we find, for $t=\tau+k$,

$$\mathsf{G}_{x}\sum_{j=0}\mathsf{J}_{k,j}\Delta\mathbb{E}_{\mathcal{E}_{\Upsilon}}\left[\hat{Y}_{\sigma\sigma,\tau+j}|\mathcal{E}_{\Upsilon}^{\tau}\right]+\mathsf{H}^{SV}_{\sigma\sigma,k}\mathcal{E}_{\Upsilon,\tau}+\mathsf{G}_{Y}\Delta\mathbb{E}_{\mathcal{E}_{\Upsilon}}\left[\hat{Y}_{\sigma\sigma,\tau+k}|\mathcal{E}_{\Upsilon}^{\tau}\right]=0$$

where $\mathsf{H}^{SV}_{\sigma\sigma,k}=\mathcal{I}^{(a)}\cdot\mathsf{B}^{SV}_{\sigma\sigma,k}+\rho_{\Upsilon}^{k}\int\mathsf{x}_{\Upsilon}d\Omega^{*}$. As this is is a linear equation that only depends on τ through $\mathcal{E}_{\Upsilon,\tau}$, this implies $\Delta\mathbb{E}_{\mathcal{E}_{\Upsilon}}\left[\hat{X}_{\sigma\sigma,\tau+k}|\mathcal{E}_{\Upsilon}^{\tau}\right]=\hat{X}^{SV}_{\sigma\sigma,k}\mathcal{E}_{\Upsilon,\tau}$ where $\hat{X}^{SV}_{\sigma\sigma,k}$ solves (54). Our knowledge of $\mathbb{E}_{\mathcal{E}_{\Upsilon}}\left[\hat{X}^{\Upsilon}_{\sigma\sigma,t}\right]$ and $\Delta\mathbb{E}_{\mathcal{E}_{\Upsilon}}\left[\hat{X}^{\Upsilon}_{\sigma\sigma,\tau+k}|\mathcal{E}_{\Upsilon}^{\tau}\right]$ immediately implies $\hat{X}^{\Upsilon}_{\sigma\sigma,t}\left(\mathcal{E}_{\Upsilon}^{t}\right)=\hat{X}_{\sigma\sigma,t}+\sum_{s=0}^{t}\hat{X}^{SV}_{\sigma\sigma,t-s}\mathcal{E}_{\Upsilon,s}$. Finally, note that $\mathsf{B}^{SV}_{\sigma\sigma,k}=\sum_{j=0}^{k-1}\rho_{\Upsilon}^{k-j+1}\left(\mathcal{L}^{(a)}\right)^{j}\cdot\mathcal{M}\cdot\mathsf{px}^{\mathsf{SV}}_{\sigma\sigma}$, which implies that $\mathsf{B}^{SV}_{\sigma\sigma,k+1}=\rho_{\Upsilon}\mathsf{B}^{SV}_{k}+\left(\mathcal{L}^{(a)}\right)^{k}\cdot\mathcal{M}\cdot\mathsf{px}^{\mathsf{SV}}_{\sigma\sigma}$ and that $\mathsf{H}^{SV}_{\sigma\sigma,k}$ therefore satisfies the recursion in Proposition 2^{HA}.

D.3 Proofs for Section 5.3

For this section we will allow $\mathcal{E}$ to be multivariate. This implies that all the directional derivatives $\hat{X}_{t}$ and $\hat{x}_{t}(a,\theta)$ should be interpreted as matrices. We let $\Sigma_{\mathcal{E}}$ represent the covariance matrix of $\mathcal{E}$.

D.3.1 Proof of Lemma 3^{PP}

We begin by differentiation equation (60) in direction $\hat{Z}_{t}$. This implies that $\mathsf{s}\mathbb{E}_{\varepsilon}\left[\hat{x}_{t+1}|a,\theta\right]\mathsf{R}\overline{Y}+\mathbb{E}_{\varepsilon}\left[\mathsf{s}\overline{x}|a,\theta\right]\mathsf{R}\hat{Y}_{t+1}=0$. In the steady state $\mathsf{R}\overline{Y}=0$, which implies that this equation can only hold if $\mathsf{R}\hat{Y}_{t}=0$ when $t\geq1$, however this places no restrictions on $\hat{R}^{x}_{0}:=\mathsf{R}\hat{Y}_{0}$.

Next we differentiate equation (59) in direction $\hat{Z}_{t}$. Doing so yields

$$\mathsf{F}_{x}(a,\theta)\hat{x}_{t}(a,\theta,k)+\mathsf{F}_{Y}(a,\theta)\hat{Y}_{t}+\mathsf{F}_{x^{e}}(a,\theta)\left(\mathbb{E}_{\varepsilon}\left[\overline{x}_{a}|a,\theta\right]\mathsf{p}\hat{x}_{t}(a,\theta,k)+\mathbb{E}_{\varepsilon}\left[\hat{x}_{t+1}|a,\theta\right]\right)+\mathsf{F}_{k}(a,\theta)k^{\mathsf{T}}\hat{R}^{x}_{t}=0. \tag{85}$$

For $t \geq 1$, $\hat{R}^x_t = 0$ implies that this is equivalent to (74) and is solved by $\hat{x}_t(a,\theta,k) = \sum_{s=0}^{\infty} \mathsf{x}_s(a,\theta)\overline{Y}_{Z,t+s}$. For $t = 0$, substitute for $\hat{x}_1(a,\theta)$ in (85) and solve for $\hat{x}_0(a,\theta,k)$ to obtain (62).

D.3.2 Proof of Lemma 7

To determine $\overline{k}(a,\theta,k)$ we differentiate (60) twice with respect to σ to get $\mathbb{E}_\varepsilon\left[\mathsf{s}\overline{x}|a,\theta\right]\overline{R}^x_{\sigma\sigma}+\mathbb{E}_\varepsilon\left[\left(s\hat{x}_0\right)\mathcal{E}\mathcal{E}^\intercal\left(\hat{R}^x_0\right)^\intercal|a,\theta\right] = 0$. As $\mathcal{E}$ is independent of θ we conclude that this is equivalent to (after substituting for $\hat{x}_0$)

$$\mathbb{E}_\varepsilon\left[\mathsf{s}\overline{x}|a,\theta\right]\overline{R}^x_{\sigma\sigma} + \sum_{s=0}^{\infty}\mathbb{E}_\varepsilon\left[\mathsf{s}\mathsf{x}_s|a,\theta\right]\hat{Y}_s\Sigma_{\mathcal{E}}\left(\hat{R}^x_0\right)^\intercal + \mathbb{E}_\varepsilon\left[\mathsf{s}\mathsf{r}|a,\theta\right]\overline{k}(a,\theta,k)^\intercal\hat{R}^x_0\Sigma_{\mathcal{E}}\left(\hat{R}^x_0\right)^\intercal = 0.$$

Define $\mathfrak{S}\left(\hat{R}^x_0\right) := \left(\hat{R}^x_0\Sigma_{\mathcal{E}}\left(\hat{R}^x_0\right)^\intercal\right)$ and exploit the knowledge that $\mathbb{E}\left[\mathsf{S}\mathsf{r}|a,\theta\right]$ is a real number to solve for

$$\overline{k}(a,\theta,k)^\intercal = \mathsf{v}_{\sigma\sigma}(a,\theta)\overline{R}^x_{\sigma\sigma}\mathfrak{S}\left(\hat{R}^x_0\right)^{-1} + \sum_s \mathsf{v}_s(a,\theta)\hat{Y}_s\Sigma_{\mathcal{E}}\left(\hat{R}^x_0\right)^\intercal\mathfrak{S}\left(\hat{R}^x_0\right)^{-1}.$$

D.3.3 Proof of Proposition 7

Differentiating the G mapping in direction $\hat{Z}_t$ to get $\mathsf{G}_Y\hat{Y}_t+\mathsf{G}_x\left(\int\overline{x}d\Omega\right)_Z\cdot\hat{Z}_t = 0$, where $\left(\int\overline{x}d\Omega\right)_Z\cdot\hat{Z}_t = \int\hat{x}_t d\Omega^* + \int\overline{x}d\hat{\Omega}_t$. As $\overline{x}(a,\theta,k) = \overline{x}(a,\theta)$ is independent of k, it suffices to know the deviation of the marginal distribution $\hat{\Omega}_t(a,\theta) = \lim_{k\to\infty}\hat{\Omega}(a,\theta,k)$. For the rest of this proof by using $\hat{\Omega}_t$ to represent this marginal distribution.

Following same steps as in the proof of Lemma 4^{HA} implies that for $t \geq 1$ $\frac{d}{d\theta}\hat{\Omega}_{t+1} = \mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_t - \mathcal{M}\cdot\hat{a}_t$, where $\mathcal{L}^{(a)}$ and $\mathcal{M}$ are as defined in the main text, and

$$\frac{d}{d\theta}\hat{\Omega}_1\langle a',\theta'\rangle = -\iiint\overline{\Lambda}(a',\theta',a,\theta)\hat{a}_0(a,\theta,k)\omega^*(a,\theta,k)dkdad\theta.$$

Substituting for $\hat{a}_0$ using Lemma 3^{PP} we find $\frac{d}{d\theta}\hat{\Omega}_1 = -\sum_s\mathcal{M}\cdot\mathsf{px}_s\hat{Y}_s - \mathcal{M}^{PP}\cdot(k^*)^\intercal\hat{R}^x_0$, where $\mathcal{M}^{PP}$ and k^* are as defined in the main text. As $\overline{k}(a,\theta)$ is independent of k it is straightforward to show that

$$\begin{aligned} k^*(a',\theta') &= \iint\overline{\Lambda}(a',\theta',a,\theta)\overline{k}(a,\theta)\omega^*(a,\theta)dad\theta = \left(\mathcal{M}\cdot\overline{k}\right)(a',\theta'). \\ &= \underbrace{\left(\mathcal{M}\cdot\mathsf{v}_{\sigma\sigma}\right)(a',\theta')}_{k^*_{\sigma\sigma}(a',\theta')}\overline{R}^x_{\sigma\sigma}\mathfrak{S}\left(\hat{R}^x_0\right)^{-1} + \sum_s\underbrace{\left(\mathcal{M}\cdot\mathsf{v}_s\right)(a',\theta')}_{k^*_s(a',\theta')}\hat{Y}_s\Sigma_{\mathcal{E}}\left(\hat{R}^x_0\right)^\intercal\mathfrak{S}\left(\hat{R}^x_0\right)^{-1}. \end{aligned}$$

We use this to show the following claim

Claim 19. $\frac{d}{d\theta}\hat{\Omega}_t$ is given by $\frac{d}{d\theta}\hat{\Omega}_t = -\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\hat{Y}_s - \left(\mathcal{L}^{(a)}\right)^{t-1}\cdot\mathcal{M}^{PP}\cdot(k^*)^\intercal\hat{R}^x_0$.

Proof. Time $t = 1$ holds trivially as $\frac{d}{d\theta}\hat{\Omega}_t$ and $\mathsf{A}_{1,s} = \mathcal{M}\cdot\mathsf{px}_s$. For $t > 1$, we proceed by induction as

$$\begin{aligned} \frac{d}{d\theta}\hat{\Omega}_{t+1} &= \mathcal{L}^{(a)}\cdot\left(-\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\hat{Y}_s - \left(\mathcal{L}^{(a)}\right)^{t-1}\cdot\mathcal{M}^{PP}\cdot(k^*)^\intercal\hat{R}^x_0\right) - \sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{px}_{s-t}\hat{Y}_s \\ &= -\sum_{s=0}^{\infty}\left(\mathcal{L}^{(a)}\cdot\mathsf{A}_{t,s} + \mathcal{M}\cdot\mathsf{px}_{s-t}\right)\hat{Y}_s - \left(\mathcal{L}^{(a)}\right)^t\cdot\mathcal{M}^{PP}\cdot(k^*)^\intercal\hat{R}^x_0 \end{aligned}$$

which completes the proof. ∎

Finally, as $\int \hat{x}_t d\Omega^* = \sum_{s=0} \int \mathsf{x}_{s-t} d\Omega^* \hat{Y}_s + \int \mathsf{r}(a,\theta)(k^*)^\intercal \, dad\theta \hat{R}_0^x$, we can conclude that

$$\left(\int \overline{x} d\Omega\right)_Z \cdot \hat{Z}_t = \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \hat{Y}_s + \mathcal{N}_t \cdot (k^*)^\intercal \hat{R}_0^x$$

where $\mathcal{N}_t$ is as defined in the main text. To complete the proof we note that in the $\sigma = 0$ limit (61) simplifies to

$$\overline{R}_{\sigma\sigma}^x \mathfrak{S}\left(\hat{R}_0^x\right)^{-1} = \frac{\overline{K}^\intercal}{\mathsf{V}_{\sigma\sigma}} - \sum_s \mathsf{V}_s \hat{Y}_s \Sigma_{\mathcal{E}} \left(\hat{R}_0^x\right)^\intercal \mathfrak{S}\left(\hat{R}_0^x\right)^{-1}$$

where $\mathsf{V}_{\sigma\sigma} = \int k_{\sigma\sigma}^* dad\theta$ and $\mathsf{V}_s = \int k_s^* dad\theta$. Thus,

$$\left(\int \overline{x} d\Omega\right)_Z \cdot \hat{Z}_t = \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \hat{Y}_s + \sum_{s=0}^{\infty} \mathsf{J}_{t,s}^{PP,1} \hat{Y}_s \Sigma_{\mathcal{E}} \left(\hat{R}_0^x\right)^\intercal \mathfrak{S}\left(\hat{R}_0^x\right)^{-1} \hat{R}_0^x + \mathsf{J}_t^{PP,2} \hat{Y}_0$$

where $\mathsf{J}_{t,s}^{PP} = \mathcal{N}_t \cdot k_s^* - \mathcal{N}_t \cdot k_{\sigma\sigma}^* \frac{\mathsf{V}_s}{\mathsf{V}_{\sigma\sigma}}$ and $\mathsf{J}_t^{PP,2} = \mathcal{N}_t \cdot k_{\sigma\sigma}^* \frac{\overline{K}^\intercal}{\mathsf{V}_{\sigma\sigma}} \mathsf{R}$, which implies

$$\mathsf{G}_Y \hat{Y}_t + \mathsf{G}_x \left(\sum_{s=0}^{\infty} \mathsf{J}_{t,s} \hat{Y}_s + \sum_{s=0}^{\infty} \mathsf{J}_{t,s}^{PP,1} \hat{Y}_s \Sigma_{\mathcal{E}} \left(\hat{R}_0^x\right)^\intercal \mathfrak{S}\left(\hat{R}_0^x\right)^{-1} \hat{R}_0^x + \mathsf{J}_t^{PP,2} \hat{Y}_0\right) = 0.$$

In the special case where Θ is 1 dimensional we have $\Sigma_{\mathcal{E}} \left(\hat{R}_0^x\right)^\intercal \mathfrak{S}\left(\hat{R}_0^x\right)^{-1} \hat{R}_0^x = 1$ and thus $\mathsf{G}_Y \hat{Y}_t + \mathsf{G}_x \sum_{s=0}^{\infty} \left(\mathsf{J}_{t,s} + \mathsf{J}_{t,s}^{PP}\right) \hat{Y}_s = 0$, where $\mathsf{J}_{t,s}^{PP} = \mathsf{J}_{t,s}^{PP,1} + 1_{s=0} \mathsf{J}_t^{PP,2}$, which completes the proof.

E Details for Section 7

E.1 Calibration

The baseline parameter values are given in the following table.

Table 2: CALIBRATION OF THE KRUSELL-SMITH ECONOMY

Parameter	Description	Value
α	Capital share	0.36
β	Discount factor	0.983
γ	Risk aversion	5
δ	Depreciation rate of capital	1.77%
ϕ	Adjustment cost of capital	35
ρ_ϵ	Idiosyncratic mean reversion	0.966
$\sigma_\epsilon/\sqrt{1-\rho_\epsilon^2}$	Cross-sectional std of log earnings	0.503
ρ_Θ	Aggregate mean reversion	0.80
σ_Θ	Std of Aggregate TFP shocks	0.014
N_ϵ	Points in Markov chain for ϵ	7
N_z	Grid points for the policy rule $\bar{x}^i(z)$	60
I_z	Grid points for the distribution $\bar{\omega}_i$	1000
T	Time horizon (in quarters) for IRF	400

E.2 Transitions across steady-states

In this section, we show how to apply our method to compute deterministic transitions across two steady states. We modify the aggregate TFP process to have a parameter $\overline{\Theta}$ that controls the mean and consider a one-time permanent change of 5% to $\overline{\Theta}$. In the economy with high TFP, the distribution of savings shifts to the right to accommodate the higher demand of the capital which is now more productive.

To apply the insights from Section 5.1, we need to compute $\hat{\Omega}_0 = \Omega^* - \Omega_0$. We set the asset distribution in the non-stochastic economy with high TFP to be Ω^* and asset distribution in the non-stochastic economy with low TFP to be Ω_0. This allows us to construct the new term in Lemma 1^{TD}, $\mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \frac{d}{d\theta}\hat{\Omega}_0$ which takes negligible amount of time given that we have precomputed operators $\mathcal{I}^{(a)}$ and $\mathcal{L}^{(a)}$. We truncate T when the difference between $\overline{X}\left(\Omega^*; \overline{\Theta} = 1\right) + \overline{X}_{Z,0}$ and $\overline{X}\left(\Omega^*; \overline{\Theta} = .95\right)$ is below a threshold.

In Figure 6, we plot the first-order expansions of the mean path of aggregate capital and the distribution of capital between the two steady states. We see capital slowly approaching a higher level and the distribution of wealth shifts rightwards.

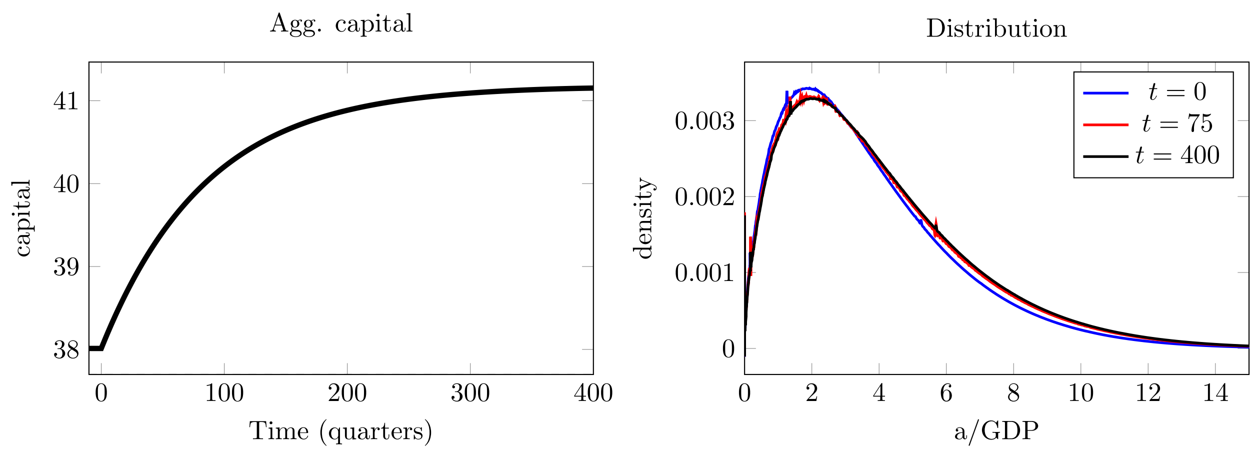
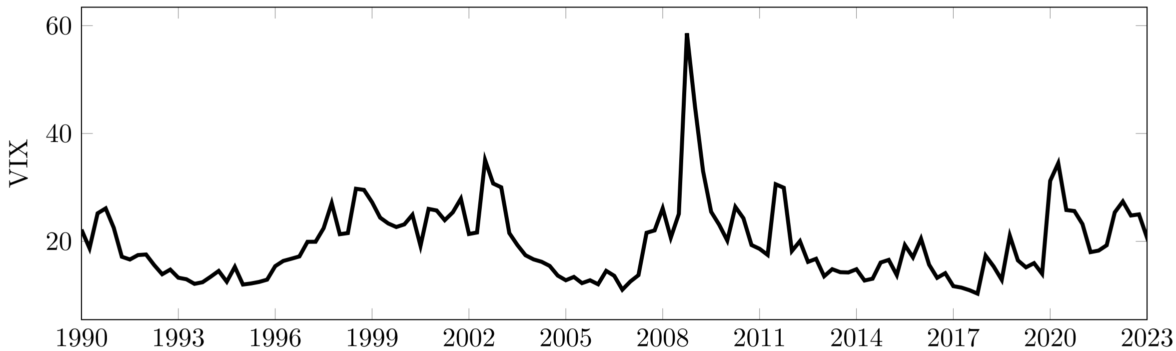


Figure 6: Transition path for capital and the distribution of savings after a 5% permanent increase in agg. TFP

E.3 More Details for Stochastic Volatility

In our baseline with a scalar aggregate shocks, the changes in volatility of all aggregate variables are proportional to $1 + \Upsilon_t$. To calibrate the stochastic process for Υ_t,we use the observed fluctuations in the CBOE Volatility Index (or the VIX). The VIX measures market expectation of near term standard deviation of stock returns conveyed by stock index option prices. The data is quarterly and the sample extends from 1990Q1 to 2023Q2. In Figure 7, we plot the VIX series.

Figure 7: AGGREGATE UNCERTAINTY



Notes: The figure plots the quarterly series of CBOE VIX for the sample period 1990Q1 to 2023Q2.

We use the approximation that $1+\Upsilon_t \approx \exp\{\Upsilon_t\}$ and use $\exp\{\Upsilon_t\} = \left[\frac{VIX_t}{(\overline{VIX})}\right]^2$ to estimate $(\rho_\Upsilon, \sigma_\Upsilon)$ by estimating the following specificiation using ordinary least squares.

$$\ln VIX_t = c + \rho_\Upsilon \ln VIX_{t-1} + \left(\frac{\sigma_\Upsilon}{2}\right)\epsilon_t \tag{86}$$

and the residual ϵ_t is assumed to be mean zero and serially uncorrelated. The following table summarizes the results.

Table 3: ESTIMATING Υ_t

Parameter	Value	Standard Error
c	0.60	0.16
ρ_Υ	0.80	0.05
σ^2_Υ	0.16	0.019

Notes: The table reports the OLS estimates of $\rho_\Upsilon$and σ^2_Υ using specification (86).

Online Supplementary Material

A Additional Numerical Details for Section 4

There are multiple approaches to approximating derivatives with respect to the individual state, e.g. $\overline{x}_a(a,\theta)$ and $\overline{x}_{aa}(a,\theta)$. The first way is direct: use the derivatives of basis functions to evaluate derivatives of $\overline{x}$, e.g. $\overrightarrow{x}_a = \overline{x}^{\#}\Phi_a$ and $\overrightarrow{x}_{aa} = \overline{x}^{\#}\Phi_{aa}$. The alternative is to compute the derivatives directly. To see how this is done we can differentiate the F mapping with respect to a to obtain

$$\mathsf{F}_a(a,\theta) + \mathsf{F}_x(a,\theta)\overline{x}_a(a,\theta) + \mathsf{F}_{x^e}\mathbb{E}_\varepsilon\left[\overline{x}_a, a, \theta\right]\mathsf{p}\overline{x}_a(a,\theta) = 0.$$

We can apply this formula to compute the values of $\overline{x}_a$ at the coarse grid-points

$$\overrightarrow{x}_a[j] = -\left(\overrightarrow{\mathsf{F}}_x[j] + \overrightarrow{\mathsf{F}}_{x^e}[j]\left(\mathsf{p}\overline{x}^{\#}\widetilde{\Phi}^e_a\right)[j]\right)^{-1}\overrightarrow{\mathsf{F}}_a[j]$$

and then recover the spline coefficients $\overline{x}^{\#}_a = \overrightarrow{x}_a\widetilde{\Phi}^{-1}$. The derivatives can alternatively be computed as $\overrightarrow{x}_a = \overline{x}^{\#}_a\Phi$ and $\overrightarrow{x}_{aa} = \overline{x}^{\#}_a\Phi_a$. In our experiments we found that all of these choices resulted in very similar aggregate responses. Using $\overline{x}^{\#}\Phi_a$ or $\overline{x}^{\#}_a\Phi$ resulted in nearly identical aggregate responses, so we opted to use $\overline{x}^{\#}\Phi_a$ for simplicity. Relative to the global solution, $\overline{x}^{\#}_a\Phi_a$ produced slightly smaller errors than $\overline{x}^{\#}\Phi_{aa}$ so we use $\overline{x}^{\#}_a\Phi_a$ and $\overline{x}^{\#}_a\widetilde{\Phi}^e_a$ when computing $\overline{x}_{aa}(a,\theta)$ and $\mathbb{E}_\varepsilon\left[\overline{x}_{aa}|a,\theta\right]$ respectively.

B Multivariate Extension

Here we extend our analysis to allow for a, θ, and Θ to be multidimensional. For the remainder of this section, we will let a^j represent the j^{th} element of a and θ^j represent the j^{th} element of θ. Almost all of the results extend directly with the caveat that the derivatives with respect to a, such as $\overline{x}_a(a,\theta)$, should now be viewed as matrices as opposed to vectors. In addition, the directions $\hat{Z}_t$ should be viewed as vectors with $\hat{Z}^j_t$ being the directions associated with the shocks Θ^j,and $\hat{\Omega}^j_t$ represents the associated change in the distribution.

B.1 First-order Approximation

Lemma 3 remains unchanged. The first difference comes with Lemma 4^{HA}. The operators $\mathcal{L}^{(a)}$, $\mathcal{M}$, and $\mathcal{I}^{(a)}$ remain essentially the same

$$\left(\mathcal{M}\cdot y\right)\langle a',\theta'\rangle := \int \overline{\Lambda}(a',\theta',a,\theta)y\left(a,\theta\right)d\Omega^*\left(a,\theta\right),$$

$$\left(\mathcal{L}^{(a)}\cdot y\right)\langle a',\theta'\rangle := \int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)y\left(a,\theta\right)dad\theta$$

$$\mathcal{I}^{(a)}\cdot \mathsf{y} := \int \overline{x}_a(a,\theta)\mathsf{y}(a,\theta)dad\theta,$$

with the understanding that now y is vector valued, while $\overline{a}_a$ and $\overline{x}_a$ are matrices. For notational simplicity we let $\frac{d}{d\theta}$ represent $\frac{\partial^{n_\theta}}{\partial\theta^1\partial\theta^2\cdots\partial\theta^{n_\theta}}$ and $\frac{d}{da}$ represent $\frac{\partial^{n_a}}{\partial a^1\partial a^2\cdots\partial a^{n_a}}$. For any vector valued function y define $\nabla_a\mathsf{y} = \sum_j \frac{\partial}{\partial a^j}\mathsf{y}^j$. With this definition we can generalize (4^{HA}) as follows

Lemma 4(FO MV). *For any t, $\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_t = \nabla_a\cdot\hat{\omega}_t$ where $\hat{\omega}_t$ satisfies a recursion*

$$\hat{\omega}_{t+1} = \mathcal{L}^{(a)}\cdot\hat{\omega}_t - \mathcal{M}\cdot\hat{a}_t \tag{87}$$

with $\hat{\omega}_0 = \mathbf{0}$.

Proof. We proceed by induction. It trivially holds for $t = 0$ as $\hat{\Omega}_0 = 0$. Assuming true for t we can differentiate the LoM in direction $\hat{Z}_t$ to get

$$\hat{\Omega}_{t+1}\langle a', \theta'\rangle = \iint \prod_{i=1}^{n_a} \iota\left(\overline{a}^i(a,\theta) \leq a'^i\right) \prod_{k=1}^{n_\theta} \iota\left((\rho_\theta\theta + \varepsilon)^k \leq \theta'^k\right) \mu(\varepsilon) d\epsilon d\hat{\Omega}_t$$
$$- \sum_{j=1}^{n_a} \iint \delta\left(\overline{a}^j(a,\theta) - a'^j\right) \prod_{i \neq j} \iota\left(\overline{a}^i(a,\theta) \leq a'^i\right) \prod_{k=1}^{n_\theta} \iota\left((\rho_\theta\theta + \varepsilon)^k \leq \theta'^k\right) \mu(\varepsilon) d\epsilon \hat{a}_t(a,\theta) d\Omega^*.$$

Applying $\frac{d}{da}\frac{d}{d\theta}$ to both sides yields

$$\begin{aligned}
\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{t+1}\langle a', \theta'\rangle &= \iint \prod_{i=1}^{n_a} \delta\left(\overline{a}^i(a,\theta) - a'^i\right) \prod_{k=1}^{n_\theta} \delta\left((\rho_\theta\theta + \varepsilon)^k - \theta'^k\right) \mu(\varepsilon) d\epsilon d\hat{\Omega}_t \\
&\quad - \sum_{j=1}^{n_a} \frac{\partial}{\partial a'^j} \iint \prod_{i=1}^{n_a} \delta\left(\overline{a}^i(a,\theta) - a'^i\right) \prod_{k=1}^{n_\theta} \delta\left((\rho_\theta\theta + \varepsilon)^k - \theta'^k\right) \mu(\varepsilon) d\epsilon \hat{a}_t^j(a,\theta)\, d\Omega^* \\
&= \int \overline{\Lambda}(a', \theta', a, \theta) d\hat{\Omega}_t - \sum_{j=1}^{n_a} \frac{\partial}{\partial a'^j} \int \overline{\Lambda}(a', \theta', a, \theta) \hat{a}_t^j(a,\theta)\, d\Omega^*
\end{aligned}$$

where in the second line we used the equality definition

$$\overline{\Lambda}(a', \theta', a, \theta) = \int \prod_{k=1}^{n_a} \delta\left(\overline{a}^k(a,\theta) - a'^k\right) \prod_{l=1}^{n_\theta} \delta\left((\rho_\theta\theta + \varepsilon)^l - \theta'^l\right) \mu(\varepsilon) d\epsilon = \prod_{k=1}^{n_a} \delta\left(\overline{a}^k(a,\theta) - a'^k\right) \mu\left(\theta' - \rho_\theta\theta\right)$$

If we apply $\frac{\partial}{\partial a^j}$ to both sides we find

$$\begin{aligned}
\frac{\partial}{\partial a^j}\overline{\Lambda}(a', \theta', a, \theta) &= \sum_i \delta'\left(\overline{a}^i(a,\theta) - a'^i\right) \prod_{k \neq i} \delta\left(\overline{a}^k(a,\theta) - a'^k\right) \mu\left(\theta' - \rho_\theta\theta\right) \overline{a}^i_{a^j}(a,\theta) \\
&= -\sum_i \frac{\partial}{\partial a'^i} \delta\left(\overline{a}^i(a,\theta) - a'^i\right) \prod_{k \neq i} \delta\left(\overline{a}^k(a,\theta) - a'^k\right) \mu\left(\theta' - \rho_\theta\theta\right) \overline{a}^i_{a^j}(a,\theta) \\
&= -\sum_i \frac{\partial}{\partial a'^i} \overline{\Lambda}(a', \theta', a, \theta) \overline{a}^i_{a^j}(a,\theta).
\end{aligned}$$

Next, we have

$$\begin{aligned}
\int \overline{\Lambda}(a', \theta', a, \theta) d\hat{\Omega}_t &= \iint \overline{\Lambda}(a', \theta', a, \theta) \frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_t(a,\theta) da d\theta \\
&= \sum_j \iint \overline{\Lambda}(a', \theta', a, \theta) \frac{\partial}{\partial a^j} \hat{\omega}_t^j(a,\theta) da d\theta \\
&= -\sum_j \iint \frac{\partial}{\partial a^j}\left(\overline{\Lambda}(a', \theta', a, \theta)\right) \hat{\omega}_t^j(a,\theta) da d\theta \\
&= \sum_i \frac{\partial}{\partial a'^i} \iint \sum_j \overline{\Lambda}(a', \theta', a, \theta) \overline{a}^i_{a^j}(a,\theta) \hat{\omega}_t^j(a,\theta) da d\theta.
\end{aligned}$$

All combined this implies that

$$\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{t+1} = \nabla_a \mathcal{L}^{(a)} \cdot \hat{\omega}_t - \nabla_a \mathcal{M} \cdot \hat{a}_t = \nabla_a \hat{\omega}_{t+1}$$

which completes the proof. $\square$

It should be noted that when the dimensionality of a is 1 Lemma 4(FO MV) is equivalent to Lemma 4^{HA} as it states that $\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_t = \frac{d}{da}\hat{\omega}_t$ and thus $\frac{d}{d\theta}\hat{\Omega}_t = \hat{\omega}_t$ which satisfies the same recursive system as in Lemma 4^{HA}. As the recursive system is identical, Corollary 1 remains unchanged since

$$\int \overline{x} d\hat{\Omega}_t = \sum_j \int \overline{x} \frac{\partial}{\partial a^j} \hat{\omega}_t^j da d\theta = -\sum_j \int \overline{x}_{a^j} \hat{\omega}_t^j da d\theta = -\mathcal{I}^{(a)} \cdot \hat{\omega}_t.$$

Finally, we have that Proposition 1^{HA} holds identically for the multivariate case with the understanding that all the derivatives with respect to Z are vector valued.

B.2 Second-Order Approximation

As with the first-order approximation, many of the Lemmas extend directly with the caveat that all derivatives with respect to a and $\hat{Z}_k$ are vector valued. For example, Lemma 5 is identical in the multivariate case.

We can extend the definitions of $\mathcal{L}^{(aa)}$, $\mathcal{L}^{(a,a)}$, $\mathcal{L}_{Z,t}$,$\mathcal{I}^{(aa)}$, and $\mathcal{I}^{(a)}_{Z,t}$ to the multivariate case

$$\begin{aligned}
\mathcal{L}^{(aa)} \cdot \mathsf{y} \langle a', \theta' \rangle &= \int \overline{\Lambda}(a', \theta', a, \theta) \overline{a}_{aa}(a, \theta) \cdot \mathsf{y}(a, \theta) da d\theta \\
\mathcal{L}^{(a,a)} \cdot \mathsf{y} \langle a', \theta' \rangle &= \int \overline{\Lambda}(a', \theta', a, \theta) \overline{a}_a(a, \theta) \mathsf{y}(a, \theta) \overline{a}_a(a, \theta)^T da d\theta \\
\mathcal{L}_{Z,t} \cdot \mathsf{y} \langle a', \theta' \rangle &= \int \overline{\Lambda}(a', \theta', a, \theta) \overline{a}_{aZ,t}(a, \theta) \mathsf{y}(a, \theta) \, da d\theta \\
\mathcal{I}^{(aa)} \cdot \mathsf{y} &= \int \overline{x}_{aa}(a, \theta) \cdot \mathsf{y}(a, \theta) da d\theta \\
\mathcal{I}^{(a)}_{Z,t} \cdot \mathsf{y} &= \int \hat{x}_{Za,t}(a, \theta) \mathsf{y}(a, \theta) da d\theta
\end{aligned}$$

with $\overline{x}_{aa}(a, \theta) \cdot \mathsf{y}(a, \theta) := \sum_{i,j} \overline{x}_{a^i a^j}(a, \theta) \mathsf{y}^{ij}(a, \theta)$ for matrix valued y. For a matrix valued function we define $\nabla_a \cdot \mathsf{y}$ as the vector

$$(\nabla_a \cdot \mathsf{y})^j (a, \theta) = \sum_i \frac{\partial}{\partial a^i} \mathsf{y}^{ij}(a, \theta),$$

which implies

$$\nabla_a^2 \cdot \mathsf{y} \equiv \nabla_a \cdot \nabla_a \cdot \mathsf{y} = \sum_{i,j} \frac{\partial}{\partial a^i} \frac{\partial}{\partial a^j} \mathsf{y}^{ij}$$

then we have the following extension of Lemma 6

Lemma 4(SO MV). *(a). For all t, k let $\hat{\omega}_{t,t+k}$ satisfy the recursion $\hat{\omega}_{0,k} = 0$,*

$$\hat{\omega}_{t,t+k} = \mathcal{L}^{(a)} \cdot \hat{\omega}_{t,t+k} - \mathcal{M} \cdot \hat{a}_{t,t+k} + \nabla_a \cdot \mathsf{c}_{t,t+k} - \mathsf{b}_{t,t+k}, \tag{88}$$

where $\mathsf{b}_{t,t+k}$ and $\mathsf{c}_{t,t+k}$ satisfy

$$\mathsf{b}_{t,t+k} = -\mathcal{L}^{(a)}_{Z,t} \cdot \frac{d}{d\theta} \hat{\Omega}_{t+k} - \mathcal{L}^{(a)}_{Z,t+k} \cdot \frac{d}{d\theta} \hat{\Omega}_t,$$

$$\mathsf{c}_{t,t+k} = \mathcal{M} \cdot \left(\hat{a}_t \odot \hat{a}_{t+k}^{\mathsf{T}}\right) - \mathcal{L}^{(a)} \cdot \left(\frac{d}{d\theta} \hat{\Omega}_t \odot \hat{a}_{t+k}^{\mathsf{T}}\right) - \mathcal{L}^{(a)} \cdot \left(\frac{d}{d\theta} \hat{\Omega}_{t+k} \odot \hat{a}_t^{\mathsf{T}}\right)^{\mathsf{T}}$$

then $\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{t,t+k} = \nabla_a \cdot \hat{\omega}_{t,t+k}$.

(b). $\hat{\omega}_{\sigma\sigma,t}$ satisfies recursion (87) *with $\hat{a}_t = \mathsf{p}\hat{x}_t$ being replaced with $\hat{a}_{\sigma\sigma,t} = \mathsf{p}\hat{x}_{\sigma\sigma,t}$ then $\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}$.*

Proof. (a) Next we take the second derivative of the LoM in direction $\hat{Z}_t$ and $\hat{Z}_k$ and adding to it the derivative of the LoM in direction $\hat{Z}_{t,k}$ yields, after applying $\frac{d}{da}\frac{d}{d\theta}$ to both sides

$$
\begin{aligned}
\frac{d}{da'}\frac{d}{d\theta'}\hat{\Omega}_{t+1,t+k+1}\langle a',\theta'\rangle &= \int \overline{\Lambda}(a',\theta',a,\theta)\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{t,t+k}\langle a,\theta\rangle dad\theta - \sum_i \frac{\partial}{\partial a'^i}\int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}^i_{t,t+k}(a,\theta)d\Omega^* \\
&+ \sum_{i,j}\frac{\partial}{\partial a'^i}\frac{\partial}{\partial a'^j}\int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}^i_t(a,\theta)\hat{a}^j_{t+k}(a,\theta)d\Omega^* \\
&- \sum_j \frac{\partial}{\partial a'^j}\int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}^j_k(a,\theta)\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle dad\theta \\
&- \sum_j \frac{\partial}{\partial a'^j}\int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}^j_t(a,\theta)\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_k\langle a,\theta\rangle dad\theta
\end{aligned}
$$

Written in vectorized form this is equivalent to

$$
\begin{aligned}
\frac{d}{da'}\frac{d}{d\theta'}\hat{\Omega}_{t+1,t+k+1}\langle a',\theta'\rangle &= \int \overline{\Lambda}(a',\theta',a,\theta)\nabla_a\cdot\hat{\omega}_{t,t+k}(a,\theta)dad\theta - \nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_{t,t+k}(a,\theta)d\Omega^* \\
&+ \nabla^2_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_t(a,\theta)\hat{a}_{t+k}(a,\theta)^\intercal d\Omega^* - \nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_{t+k}(a,\theta)\nabla_a\cdot\hat{\omega}_t(a,\theta)dad\theta \\
&- \nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_t(a,\theta)\nabla_a\cdot\hat{\omega}_{t+k}(a,\theta)dad\theta.
\end{aligned}
$$

Next, we note that for vector valued generalized functions w and y,

$$
\begin{aligned}
\int \overline{\Lambda}(a',\theta',a,\theta)\mathsf{w}(a,\theta)\nabla_a\cdot\mathsf{y}(a,\theta)d\theta da &= \sum_j\int \overline{\Lambda}(a',\theta',a,\theta)\mathsf{w}(a,\theta)\frac{\partial}{\partial a^j}\mathsf{y}^j(a,\theta)dad\theta \\
&= -\sum_j\int \overline{\Lambda}(a',\theta',a,\theta)\frac{\partial}{\partial a^j}\mathsf{w}(a,\theta)\mathsf{y}^j(a,\theta)dad\theta \\
&\quad -\sum_j\int \frac{\partial}{\partial a^j}\left(\overline{\Lambda}(a',\theta',a,\theta)\right)\mathsf{w}(a,\theta)\mathsf{y}^j(a,\theta)dad\theta \\
&= -\int \overline{\Lambda}(a',\theta',a,\theta)\mathsf{w}_a(a,\theta)\mathsf{y}(a,\theta)dad\theta \\
&\quad +\sum_{i,j}\frac{\partial}{\partial a'^i}\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}^i_{a^j}(a,\theta)\mathsf{w}(a,\theta)\mathsf{y}^j(a,\theta)dad\theta \\
&= -\int \overline{\Lambda}(a',\theta',a,\theta)\mathsf{w}_a(a,\theta)\mathsf{y}(a,\theta)dad\theta \\
&\quad +\nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\mathsf{y}(a,\theta)\mathsf{w}(a,\theta)^\intercal dad\theta.
\end{aligned}
$$

Applying this relationship implies

$$
\begin{aligned}
\frac{d}{da'}\frac{d}{d\theta'}\hat{\Omega}_{t+1,t+k+1}\langle a',\theta'\rangle &= \nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\hat{\omega}_{t,t+k}dad\theta - \nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_{t,t+k}(a,\theta)d\Omega^* \\
&+ \nabla^2_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_t(a,\theta)\hat{a}_{t+k}(a,\theta)^\intercal d\Omega^* + \nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_{Za,k}(a,\theta)\hat{\omega}_t(a,\theta)dad\theta \\
&- \nabla^2_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\hat{\omega}_t(a,\theta)\hat{a}_{t+k}(a,\theta)^\intercal dad\theta + \nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_{Za,t}(a,\theta)\hat{\omega}_k(a,\theta)dad\theta \\
&- \nabla^2_a\cdot\left(\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\hat{\omega}_k(a,\theta)\hat{a}_t(a,\theta)^\intercal dad\theta\right)^T
\end{aligned}
$$

which implies that $\frac{d}{da'}\frac{d}{d\theta'}\hat{\Omega}_{t+1,k+1} = \nabla_a\cdot\hat{\omega}_{t+1,k+1}$ where $\hat{\omega}_{t,k}$ satisfies (87).

(b) We proceed by induction which holds trivially for $t = 0$. Differentiating the LoM twice with respect to σ, adding the derivative in direction $\hat{Z}_{\sigma\sigma,t}$, and then applying $\frac{d}{da}\frac{d}{d\theta}$ to both sides yields

$$\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t+1}\langle a',\theta'\rangle = \int \overline{\Lambda}(a',\theta',a,\theta)\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}\langle a,\theta\rangle da d\theta - \nabla_a \cdot \int \overline{\Lambda}(a',\theta',a,\theta)\hat{a}_{\sigma\sigma,t}(a,\theta)d\Omega^*.$$

Using the same steps as the proof of Lemma 4(FO MV) we have

$$\int \overline{\Lambda}(a',\theta',a,\theta)\nabla_a \cdot \hat{\omega}_{\sigma\sigma,t}(a,\theta)dad\theta = \nabla_a \cdot \int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\hat{\omega}_{\sigma\sigma,t}(a,\theta)dad\theta.$$

which implies

$$\frac{d}{da'}\frac{d}{d\theta'}\hat{\Omega}_{\sigma\sigma,t+1} = \nabla_a \cdot \left(\mathcal{L}^{(a)} \cdot \hat{\omega}_t - \mathcal{M} \cdot \hat{a}_{\sigma\sigma,t}\right) = \nabla_a \cdot \hat{\omega}_{\sigma\sigma,t+1}$$

$\square$

Next we extend Corollary 2 to the multidimensional case as follows. .

Corollary 2^{MV}. *(a). For all t, k*

$$\left(\int \overline{x}d\Omega\right)_{ZZ} \cdot \hat{Z}_{t,t+k} + \left(\int \overline{x}d\Omega\right)_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right) = \sum_{s=0}^{\infty} \mathsf{J}_{t,s}\hat{Y}_{t+s,t+k+s} + \mathsf{H}_{t,t+k},$$

where $\{\mathsf{H}_{t,t+k}\}_{t,k}$ is characterized by the following linear recursive system

$$\mathsf{H}_{t,t+k} = \int \mathsf{x}_{t,t+k}d\Omega^* + \mathcal{I}^{(a)} \cdot \mathsf{B}_{t,t+k} + \mathcal{I}^{(aa)} \cdot \mathsf{C}_{t,t+k} - \mathcal{I}^{(a)}_{Z,t+k} \cdot \hat{\omega}_t - \mathcal{I}^{(a)}_{Z,t} \cdot \hat{\omega}_{t+k},$$

$$\mathsf{C}_{t+1,t+k+1} = \mathcal{M} \cdot \left(\hat{a}_t \odot \hat{a}^{\mathsf{T}}_{t+k}\right) - \mathcal{L}^{(a)} \cdot \left(\frac{d}{d\theta}\hat{\Omega}_t \odot \hat{a}^{\mathsf{T}}_{t+k}\right) - \mathcal{L}^{(a)} \cdot \left(\frac{d}{d\theta}\hat{\Omega}_{t+k} \odot \hat{a}^{\mathsf{T}}_t\right)^{\mathsf{T}} + \mathcal{L}^{(a,a)} \cdot \mathsf{C}_{t,t+k},$$

$$\mathsf{B}_{t+1,t+k+1} = \mathcal{M} \cdot \mathsf{px}_{t,t+k} - \mathcal{L}^{(a)}_{Z,t} \cdot \hat{\omega}_{t+k} - \mathcal{L}^{(a)}_{Z,t+k} \cdot \hat{\omega}_t + \mathcal{L}^{(a)} \cdot \mathsf{B}_{t,t+k} + \mathcal{L}^{(aa)} \cdot \mathsf{C}_{t,t+k}.$$

(b). For all t,

$$\int \overline{x}_{\sigma\sigma}d\Omega^* + \left(\int \overline{x}d\Omega\right)_Z \cdot \hat{Z}_{\sigma\sigma,t} = \sum_{s=0}^{\infty} \mathsf{J}_{t,s}\hat{Y}_{\sigma\sigma,s} + \mathsf{H}_{\sigma\sigma,t}$$

where $\{\mathsf{H}_{\sigma\sigma,t}\}_t$ satisfies recursion $\mathsf{H}_{\sigma\sigma,0} = \int \mathsf{x}_{\sigma\sigma}d\Omega^*$ and $\mathsf{H}_{\sigma\sigma,t} = \mathsf{H}_{\sigma\sigma,t-1} + \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M} \cdot \mathsf{px}_{\sigma\sigma}$.

Proof. (a) Start by noting

$$\left(\int \overline{x}d\Omega\right)_{ZZ} \cdot \hat{Z}_{t,t+k} + \left(\int \overline{x}d\Omega\right)_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right) = \int \hat{x}_{t,k}d\Omega^* + \int \hat{x}_t d\hat{\Omega}_{t+k} + \int \hat{x}_{t+k}d\hat{\Omega}_t + \int \overline{x}d\hat{\Omega}_{t,t+k}$$

We start with $\int \overline{x} d\hat{\Omega}_{t,t+k}$. For any matrix valued generalized function c

$$
\begin{aligned}
\mathcal{L} \cdot \nabla_a \cdot \mathsf{c}(a', \theta') &= \sum_{i,j} \int \overline{\Lambda}(a', \theta', a, \theta) \overline{a}_{a^j}(a, \theta) \frac{\partial}{\partial a^i} \mathsf{c}^{ij}(a, \theta) da d\theta \\
&= -\sum_{i,j} \int \overline{\Lambda}(a', \theta', a, \theta) \overline{a}_{a^i a^j}(a, \theta) \mathsf{c}^{ij}(a, \theta) da d\theta \\
&\quad - \sum_{i,j} \int \frac{\partial}{\partial a^i} \left(\overline{\Lambda}(a', \theta', a, \theta) \right) \overline{a}_{a^j}(a, \theta) \mathsf{c}^{ij}(a, \theta) da d\theta \\
&= -\mathcal{L}^{(aa)} \cdot \mathsf{c}(a', \theta') + \sum_{i,j,k} \frac{\partial}{\partial a'^k} \int \overline{\Lambda}(a', \theta', a, \theta) \overline{a}^k_{a^i} \overline{a}_{a^j}(a, \theta) \mathsf{c}^{ij}(a, \theta) da d\theta \\
&= -\mathcal{L}^{(aa)} \cdot \mathsf{c}(a', \theta') + \nabla_a \cdot \mathcal{L}^{(a,a)} \cdot \mathsf{c}(a', \theta').
\end{aligned}
$$

We can then proceed by induction. For $t = 0$ we have

$$
\hat{\omega}_{0,k} = -\sum_{s=0}^{\infty} \mathsf{A}_{0,s} \overline{Y}_{ZZ,s,k-t+s} - \mathsf{B}_{0,k-t} + \nabla_a \cdot \mathsf{C}_{0,k-t}
$$

since all terms are 0. If it holds for t then

$$
\begin{aligned}
\hat{\omega}_{t+1,t+k+1} &= \mathcal{L}^{(a)} \cdot \hat{\omega}_{t,t+k} - \sum_{s=0}^{\infty} \mathcal{M} \cdot \mathsf{px}_s \hat{Y}_{t+s,t+k+s} - \mathsf{b}_{t,t+k} + \nabla_a \cdot \mathsf{c}_{t,t+k} \\
&= \mathcal{L}^{(a)} \cdot \left(-\sum_{s=0}^{\infty} \mathsf{A}_{t,s} \hat{Y}_{s,k+s} - \mathsf{B}_{t,t+k} + \nabla_a \cdot \mathsf{C}_{t,t+k} \right) - \sum_{s=0}^{\infty} \mathcal{M} \cdot \mathsf{px}_{s-t} \hat{Y}_{s,k+s} \\
&\quad - \mathsf{b}_{t,t+k} + \nabla_a \cdot \mathsf{c}_{t,t+k} \\
&= -\sum_{s=0}^{\infty} \left(\mathcal{L}^{(a)} \cdot \mathsf{A}_{t,s} + \mathsf{a}_{s-t} \right) \hat{Y}_{s,k+s} - \left(\mathcal{L}^{(a)} \cdot \mathsf{B}_{t,t+k} + \mathsf{b}_{t,t+k} \right) + \mathcal{L}^{(a)} \cdot \nabla_a \cdot \mathsf{C}_{t,t+k} + \nabla_a \cdot \mathsf{c}_{t,t+k} \\
&= -\sum_{s=0}^{\infty} \mathsf{A}_{t+1,s} \hat{Y}_{s,k+s} - \left(\mathcal{L}^{(a)} \cdot \mathsf{B}_{t,t+k} + \mathcal{L}^{(aa)} \cdot \mathsf{C}_{t,t+k} + \mathsf{b}_{t,t+k} \right) + \nabla_a \cdot \left(\mathcal{L}^{(a,a)} \cdot \mathsf{C}_{t,t+k} + \mathsf{c}_{t,t+k} \right) \\
&= -\sum_{s=0}^{\infty} \mathsf{A}_{t+1,s} \hat{Y}_{s,k+s} - \mathsf{B}_{t+1,t+k+1} + \nabla_a \cdot \mathsf{C}_{t+1,t+k+1}.
\end{aligned}
$$

Finally, we have that

$$
\begin{aligned}
\int \overline{x} d\hat{\Omega}_{t,t+k} &= \int \overline{x} \nabla_a \cdot \hat{\omega}_{t,t+k} da d\theta \\
&= -\int \overline{x}_a \hat{\omega}_{t,t+k} da d\theta \\
&= -\int \overline{x}_a \left(-\sum_{s=0}^{\infty} \mathsf{A}_{t,s} \hat{Y}_{s,k+s} - \mathsf{B}_{t,t+k} + \nabla_a \cdot \mathsf{C}_{t,t+k} \right) da d\theta \\
&= \sum_{s=0}^{\infty} \left(\mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s} \right) \hat{Y}_{s,k+s} + \mathcal{I}^{(a)} \cdot \mathsf{B}_{t,t+k} + \int \overline{x}_{aa} \cdot \mathsf{C}_{t,t+k} da d\theta \\
&= \sum_{s=0}^{\infty} \left(\mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s} \right) \hat{Y}_{s,k+s} + \mathcal{I}^{(a)} \cdot \mathsf{B}_{t,t+k} + \mathcal{I}^{(aa)} \cdot \mathsf{C}_{t,t+k}
\end{aligned}
$$

as desired. Next

$$\int \hat{x}_t d\hat{\Omega}_{t+k} = \int \hat{x}_t(a,\theta)\nabla_a \cdot \hat{\omega}_{t+k}(a,\theta)dad\theta = -\int \hat{x}_{Za,t}(a,\theta)\hat{\omega}_{t+k}(a,\theta)dad\theta = \mathcal{I}^{(a)}_{Z,t} \cdot \hat{\omega}_{t+k}$$

and similarly for $\int \hat{x}_{t+k} d\hat{\Omega}_t$. Finally

$$\int \hat{x}_{t,t+k} d\Omega^* = \sum_{s=0}^{\infty} \int \mathsf{x}_{t-s} d\Omega^* \hat{Y}_{s,k+s} + \int \mathsf{x}_{t,t+k} d\Omega^*.$$

Adding the terms of $\left(\int \overline{x} d\Omega\right)_{ZZ} \cdot \hat{Z}_{t,t+k} + \left(\int \overline{x} d\Omega\right)_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_{t+k}\right)$ together completes the proof.

(b)Begin with

$$\int \overline{x}_{\sigma\sigma} d\Omega^* + \left(\int \overline{x} d\Omega\right)_Z \cdot \hat{Z}_{\sigma\sigma,t} = \int \hat{x}_{\sigma\sigma,t} d\Omega^* + \int \overline{x} d\hat{\Omega}_{\sigma\sigma,t}$$

Starting with $\int \overline{x} d\hat{\Omega}_{\sigma\sigma,t}$, as $\hat{\omega}_{\sigma\sigma,0} = 0$ we can roll forward

$$\hat{\omega}_{\sigma\sigma,t+1} = \mathcal{L}^{(a)} \cdot \hat{\omega}_{\sigma\sigma,t} - \sum_{s=0}^{\infty} \mathcal{M} \cdot \mathsf{px}_{s-t} \hat{Y}_{\sigma\sigma,s} - \mathcal{M} \cdot \mathsf{px}_{\sigma\sigma}$$

to obtain

$$\hat{\omega}_{\sigma\sigma,t} = -\sum_{s=0}^{\infty} \mathsf{A}_{t,s} \hat{Y}_{\sigma\sigma,s} - \mathsf{B}_{\sigma\sigma,t}$$

where $\mathsf{B}_{\sigma\sigma,0} = 0$ and $\mathsf{B}_{\sigma\sigma,t+1} = \mathcal{L}^{(a)} \cdot \mathsf{B}_{\sigma\sigma,t} + \mathcal{M} \cdot \mathsf{px}_{\sigma\sigma}$. Using integration by parts implies

$$\int \overline{x} d\hat{\Omega}_{\sigma\sigma,t} = \int \overline{x} \nabla_a \cdot \hat{\omega}_{\sigma\sigma,t} dad\theta = -\int \overline{x}_a \hat{\omega}_{\sigma\sigma,t} dad\theta = -\mathcal{I}^{(a)} \cdot \hat{\omega}_{\sigma\sigma,t}.$$

Adding to it

$$\int \hat{x}_{\sigma\sigma,t} d\Omega^* = \sum_{s=0}^{\infty} \int \mathsf{x}_{t-s} d\Omega^* \hat{Y}_{\sigma\sigma,s} + \int \mathsf{x}_{\sigma\sigma} d\Omega^*$$

we have $\mathsf{H}_{\sigma\sigma,t} = \int \mathsf{x}_{\sigma\sigma} d\Omega^* + \mathcal{I}^{(a)} \cdot \mathsf{B}_{\sigma\sigma,t}$. This is the same formula as Corollary 2 so $\mathsf{H}_{\sigma\sigma,t}$ satisfies the recursion $\mathsf{H}_{\sigma\sigma,t} = \mathsf{H}_{\sigma\sigma,t-1} + \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M} \cdot \mathsf{px}_{\sigma\sigma}$. □

Combining all of these insights implies that Proposition 2^{HA} remains unchanged in the multivariate extension.

B.3 Krusell and Smith with adjustment costs

Household problem Households hold shares in a mutual fund and date t holdings of i are denoted by $a_{i,t}$. Let D_t and P_t be the time t dividend and the price per share of the mutual fund. The household problem is given by

$$\max_{c_{i,t}, i_{i,t}, k_{i,t}} \mathbb{E} \sum_t \beta^t U(c_{i,t})$$

subject to

$$c_{i,t} + P_t a_{i,t} = W_t \exp\{\theta_{i,t}\} + (D_t + P_t) a_{i,t-1}$$

$$a_{i,t} \geq 0$$

The Euler equation of the household is given by

$$1 = \mathbb{E}_t \left(\frac{\beta U_c \left(c_{i,t+1} \right)}{U_c \left(c_{i,t} \right) + \zeta_{i,t}} \right) \left(\frac{D_{t+1} + P_{t+1}}{P_t} \right) \tag{89}$$

where $\lambda_{i,t},\ \zeta_{i,t} \geq 0$ are the Lagrange multiplier on the sequential budget and borrowing constraint respectively.

Stochastic Discount Factors Define a process $\{m_{i,t}\}$ with $m_{i,0} = 1$ and $\frac{m_{i,t+1}}{m_{i,t}} \equiv \frac{\beta U_c(c_{i,t+1})}{U_c(c_{i,t})+\zeta_{i,t}}$. For any positive process $\{o_{i,t}\}$ define M_t with $M_0 = 1$ and $\frac{M_{t+1}}{M_t} = \int o_{i,t} \frac{m_{i,t+1}}{m_{i,t}} di$. Then aggregating (89) we get that the value of the mutual fund satisfies

$$P_t = \mathbb{E}_t \sum_j \frac{M_{t+j}}{M_t} D_{t+j}.$$

Firms Problem Firms rent captal and hire workers on a spot market to produce using a Cobb Douglas production function. Let K_t be the capital used in the production at date t, their optimality gives us

$$R_t^k = \exp\left(\Theta_t\right) K_t^{\alpha-1}$$
$$W_t = (1-\alpha) \exp\left(\Theta_t\right) K_t^{\alpha},$$

Mutual Fund Problem The mutual fund owns physical capital, makes investments subject to quadratic adjustment costs, rents out the capital to the corporate sector, and maximizes present value of dividends. For a given $\{M_t\}$, the problem of the mutual fund is

$$\max_{K_t, D_t} \mathbb{E}_0 \sum_t M_t D_t$$

$$D_t = R_t^k K_t - I_t - \frac{\phi}{2} \left(\frac{I_t}{K_t} - \delta \right)^2 K_t$$
$$K_{t+1} = (1-\delta) K_t + I_t$$

Let Q_t be the multiplier on the capital accumulation equation. The optimality of the mutual fund with respect to I_t

$$Q_t = 1 + \phi \left(\frac{I_t}{K_t} - \delta \right)$$

and with respect to K_t is

$$\mathbb{E}_t \frac{M_{t+1}}{M_t} \left\{ \frac{R_{t+1}^k + \phi \left(\frac{I_{t+1}}{K_{t+1}} - \delta \right) \frac{I_{t+1}}{K_{t+1}} - \frac{\phi}{2} \left(\frac{I_{t+1}}{K_{t+1}} - \delta \right)^2 + (1-\delta) Q_{t+1}}{Q_t} \right\} = 1. \tag{90}$$

Its easy to check that

$$\frac{R_{t+1}^k + \phi \left(\frac{I_{t+1}}{K_{t+1}} - \delta \right) \frac{I_{t+1}}{K_{t+1}} - \frac{\phi}{2} \left(\frac{I_{t+1}}{K_{t+1}} - \delta \right)^2 + (1-\delta) Q_{t+1}}{Q_t} = \frac{D_{t+1} + Q_{t+1} K_{t+2}}{Q_t K_{t+1}}$$

and thus iterating on (90) we get

$$Q_t K_{t+1} = P_t = \mathbb{E}_t \sum_j \frac{M_{t+j}}{M_t} D_{t+j}.$$

Equilibrium The equilibrium is given by

$$c_{i,t} + k_{i,t} = W_t \exp\{\theta_{i,t}\} + R_t k_{i,t-1} \tag{91a}$$

$$\lambda_{i,t} = \mathbb{E}_t \beta U_c(c_{i,t}) R_t \tag{91b}$$

$$U_c(c_{i,t}) + \zeta_{i,t} = \mathbb{E}_t \lambda_{i,t+1} \tag{91c}$$

$$k_{i,t}\zeta_{i,t} = 0 \tag{91d}$$

$$W_t - (1-\alpha)\exp(\Theta_t) K_t^{\alpha} = 0, \tag{91e}$$

$$R_t = \frac{(1-\alpha)\exp(\Theta_t) K_t^{\alpha} + \phi\left(\frac{I_t}{K_t} - \delta\right)\frac{I_t}{K_t} - \frac{\phi}{2}\left(\frac{I_t}{K_t} - \delta\right)^2 + (1-\delta)Q_t}{Q_{t-1}} \tag{91f}$$

$$K_{t+1} = (1-\delta) K_t + I_t \tag{91g}$$

$$Q_t = 1 + \phi\left(\frac{I_t}{K_{t-1}} - \delta\right), \tag{91h}$$

$$\int k_{i,t} di = Q_t K_{t+1}. \tag{91i}$$

To map the problem to Section 3.2 notation, use the following definitions

$$A_{t-1} = [Q_{t-1}, K_t]^{\intercal}$$

$$X_t = [Q_t, K_t, W_t, R_t, I_t]^{\intercal}$$

$$a_{i,t-1} = [\theta_{i,t-1}, k_{i,t-1}]^{\intercal}$$

$$x_{i,t} = [k_{i,t}, c_{i,t}, \lambda_{i,t}, \zeta_{i,t}]^{\intercal}$$